(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 670 374 A1**

(12) # EUROPEAN PATENT APPLICATION

(21) Application number: **95250047.8**

(22) Date of filing: **28.02.95**

(51) Int. Cl.6: **C12Q 1/68**, G01N 33/52, C07H 21/00, G01N 27/447

(30) Priority: **01.03.94 US 204627**

(43) Date of publication of application: **06.09.95 Bulletin 95/36**

(84) Designated Contracting States: **DE FR GB SE**

(71) Applicant: **LI-COR, INC.**
4421 Superior Street
Lincoln,
Nebraska 68504 (US)

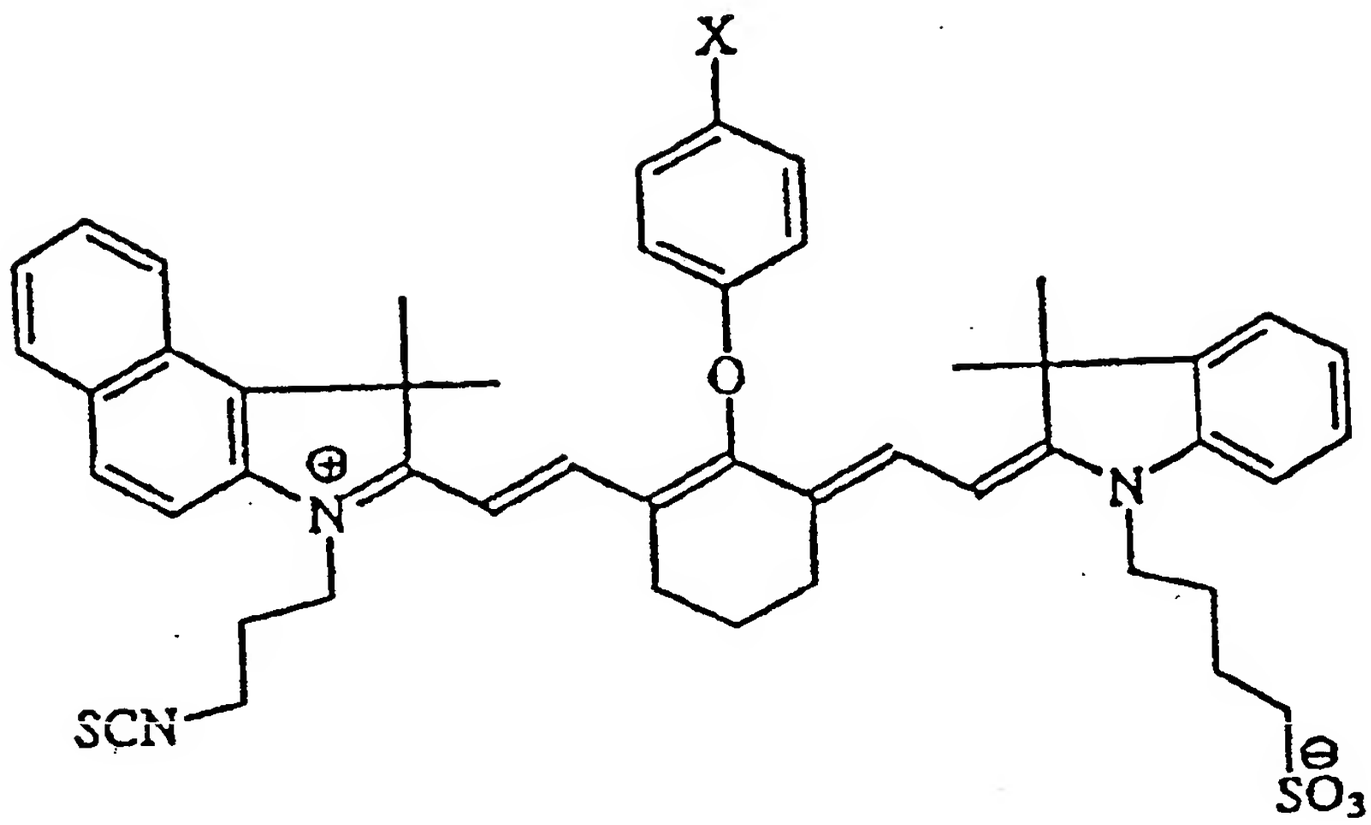
(72) Inventor: **Patonay, Gabor**
5607 Boggs Drive
Stone Mountain,
Georgia 30082 (US)
Inventor: **Narayanan, Narasimhachari**
2841 Tierra Drive No.201
Lincoln,
Nebraska 68516 (US)
Inventor: **Strekowski, Lucjan**
4275 Old Coach Court
Stone Mountain,
Georgia 30083 (US)
Inventor: **Middendorf, Lyle Richard**
8135 Sanborn Drive
Lincoln,
Nebraska 68505 (US)
Inventor: **Lipowska, Malgorzata**
120 Northern Avenue
Decatur,
Georgia 30030 (US)

(74) Representative: **UEXKÜLL & STOLBERG**
Patentanwälte
Beselerstrasse 4
D-22607 Hamburg (DE)

(54) **A method of identifying strands of DNA using infrared fluorescent labels.**

(57) The invention relates to a method of identifying strands of DNA comprising the steps of: marking the strands with fluorescent labels, irradiating the strands and detecting the light emitted from the fluorescent strands. The labels emit light in a region of wavelengths including at least one wavelength within the infrared and near infrared region wherein the fluorescent label includes a chromophore having e.g. the formula:

EP 0 670 374 A1


X
O
N
N
SCN
Ie
SO3

This invention relates to the sequencing of fluorescence labeled DNA and the detecting of the DNA after irradiation by light from a laser, and suitable labels therefor.

In one class of techniques for sequencing DNA, identical strands of DNA are marked with fluorescent dye. The strands are marked by attaching specially synthesized fluorescent oligonucleotide primers or probes to the strands of DNA or by attaching the fluorescent dye directly to the strands.

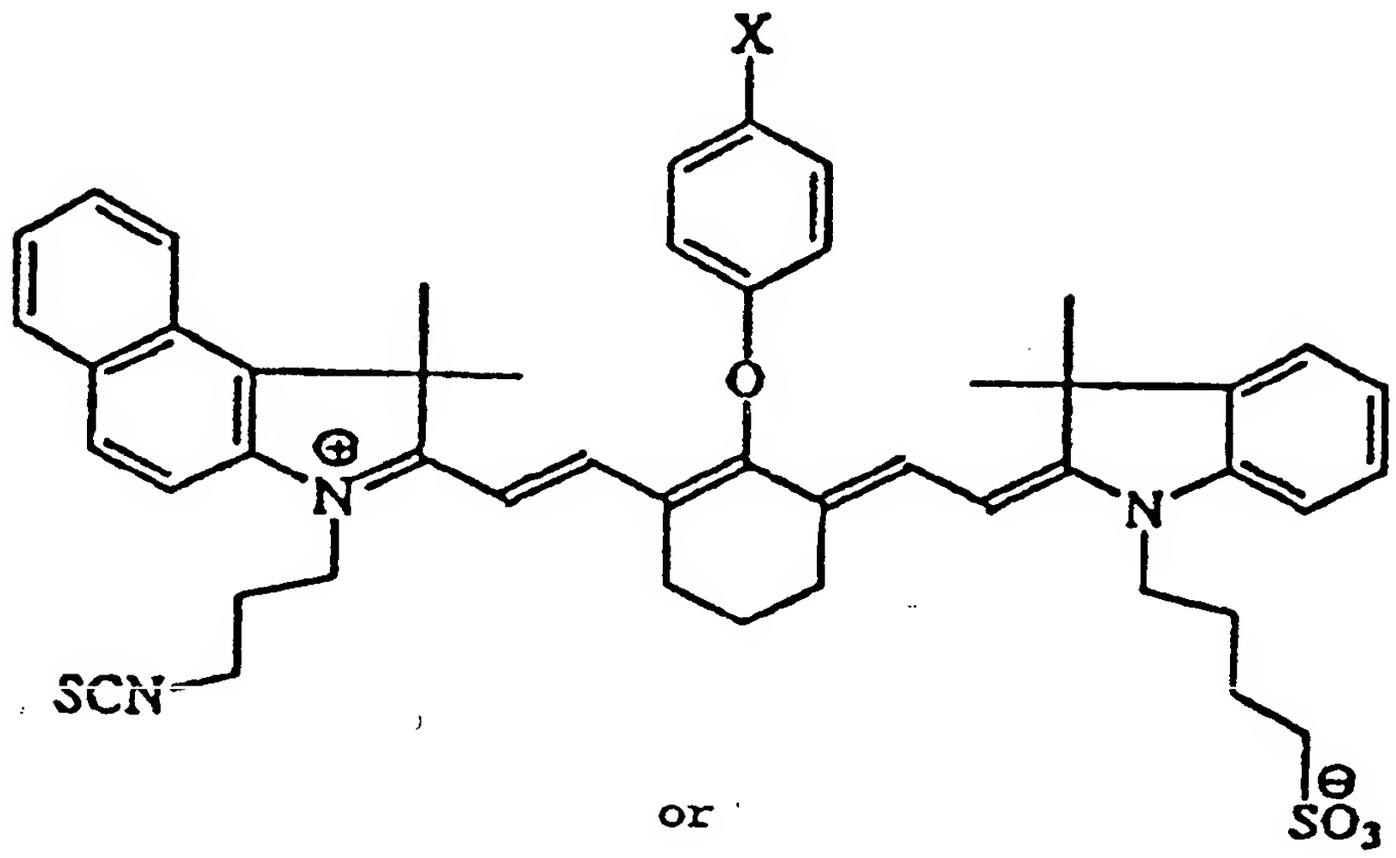
The strands are separated into four aliquots. The strands in a given aliquot are either individually cleaved at or synthesized to any base belonging to only one of the four base types, which are adenine, guanine, cytosine and thymine (hereinafter A, G, C and T). The adenine-, guanine-, cytosine- and thymine-terminated strands are then electrophoresed for separation and the separated strands are irradiated by a laser and the emission detected. The rate of electrophoresis indicates the DNA sequence.

In a prior art sequencing technique of this class, the fluorescent dyes used as markers have their maximum emission spectra in the visible range, the DNA is subject to irradiation in the visible spectra, and visible spectra detectors and light sources are used. Generally photomultipliers tubes are used for detection.

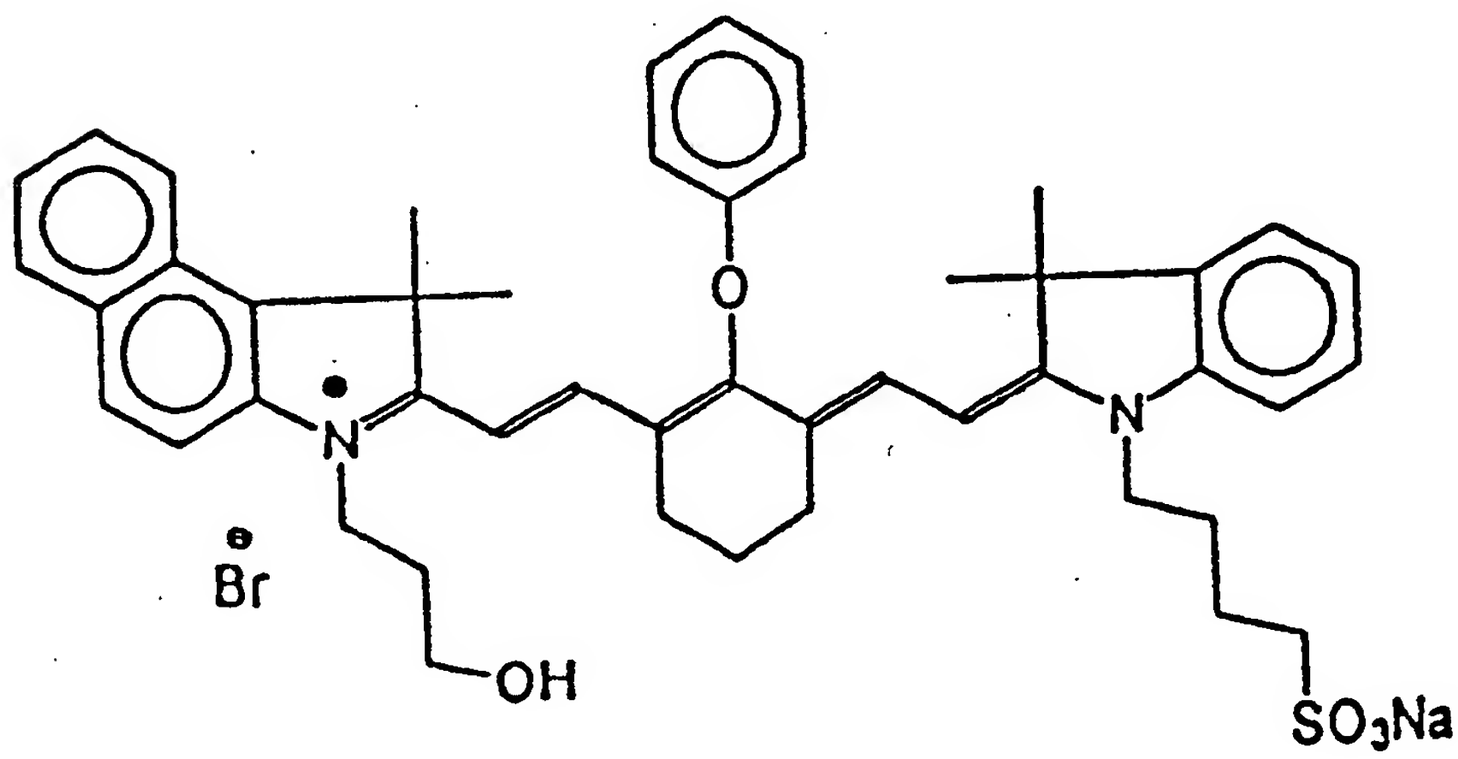
The prior art techniques for DNA sequencing have several disadvantages such as: (1) because the dyes have their emission spectra in the visible region of light spectrum, the lasers used to excite the fluorescent markers, and under some circumstances, the detectors for the light tend to be expensive; and (2) they are relatively noisy due to the background interferences by biomolecules.

Cyanine dyes are known to absorb infrared and near infrared light and techniques for the synthesis of derivatives of the cyanine dyes are known. However, obtaining absorbance and emission bands that reduce the effect of background noise during gel electrophoresis apparatus when irradiating with a diode laser scanning has been difficult.

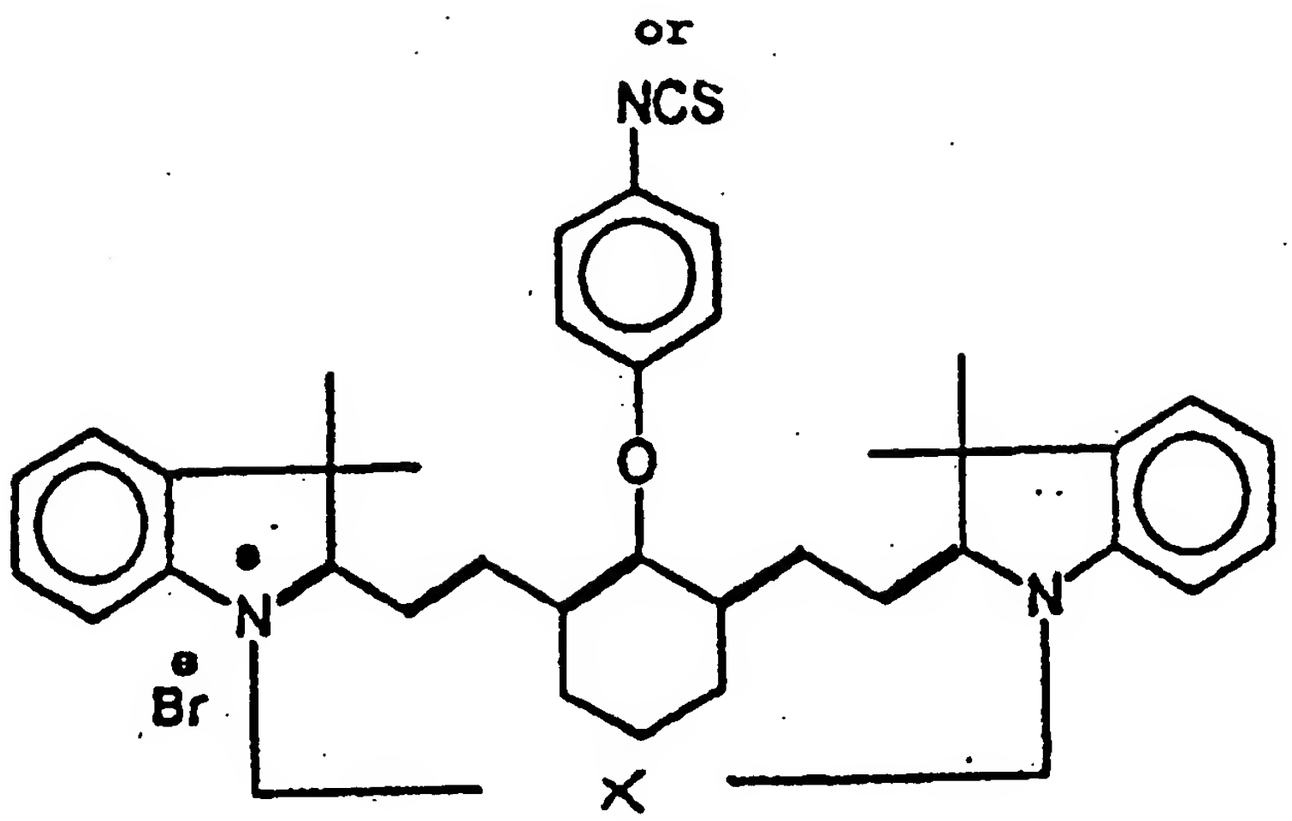
To accomplish this task, strands of DNA are marked with fluorescent labels that emit light in a region of wavelengths including at least one wavelength within the infrared and near infrared region wherein the fluorescent label includes a chromophore having one of the formulas:



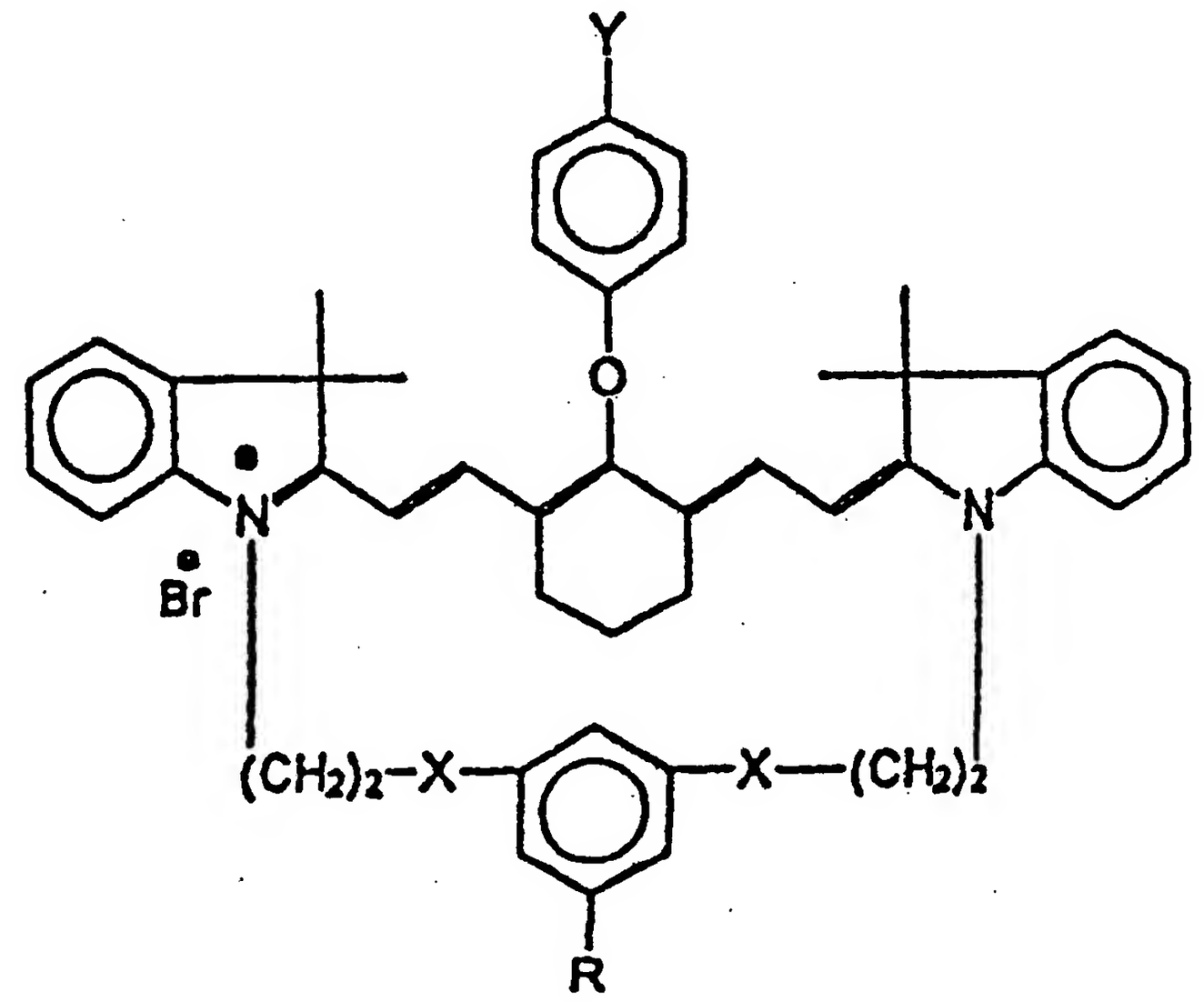
or



or



where X is $(CH_2)_n$; n = 4-10 or X is $-CH_2 - CH_2 - O - CH_2 - CH_2 - O - CH_2 - CH_2 -$;
or



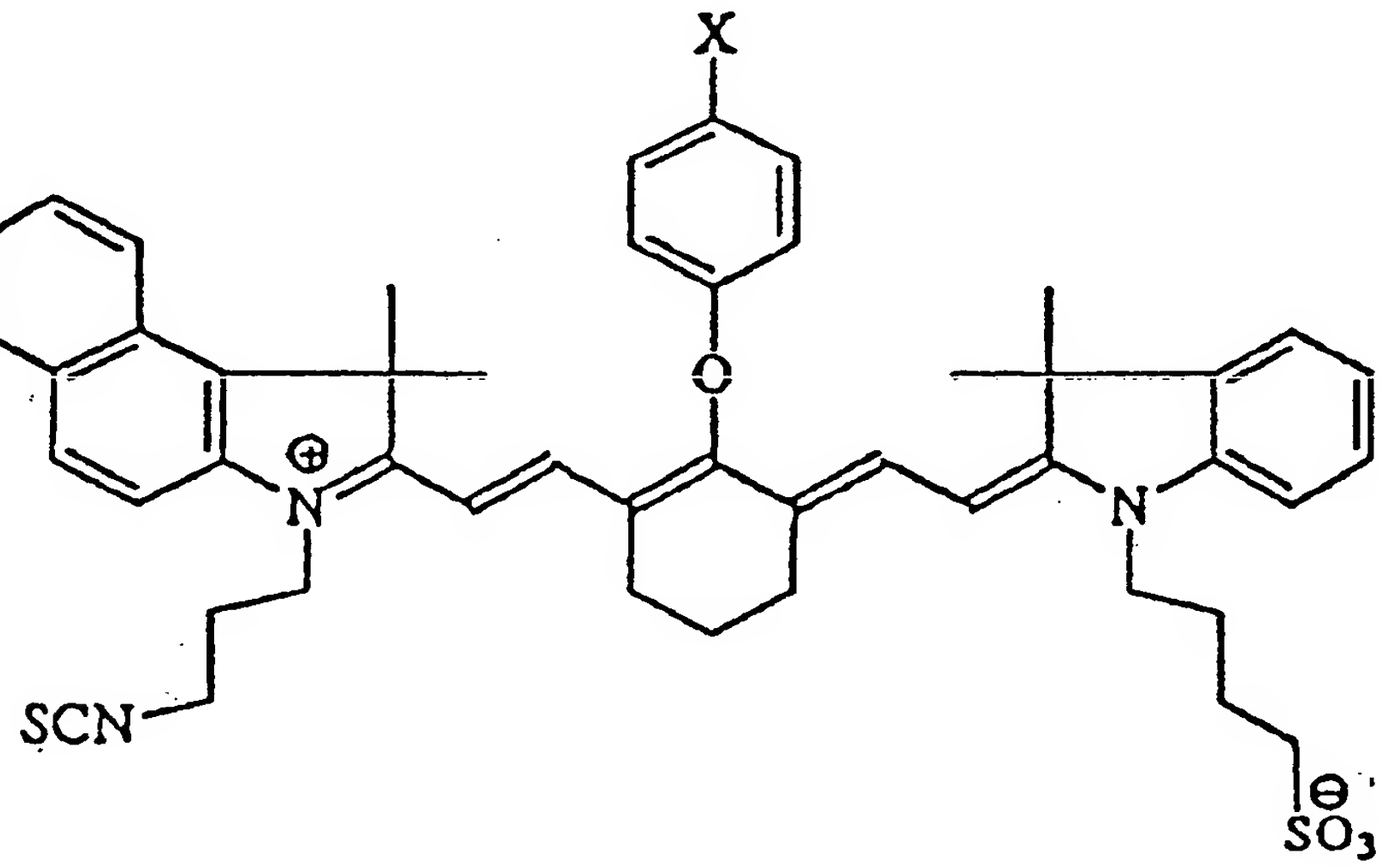
wherein X consists of one of O or NH, Y consists of one of NCS or H; and R consists of one of H, NCS, CH_2OH, CH_2NCS, COOH.

The strands are irradiated with light having a wavelength within one of the infrared and near infrared regions; and the light emitted from the fluorescent labels is detected. Advantageously, fluorescent markers are attached to probes and the probes are hybridized to the strands. The labels are detected by scanning with an infrared laser diode light source.

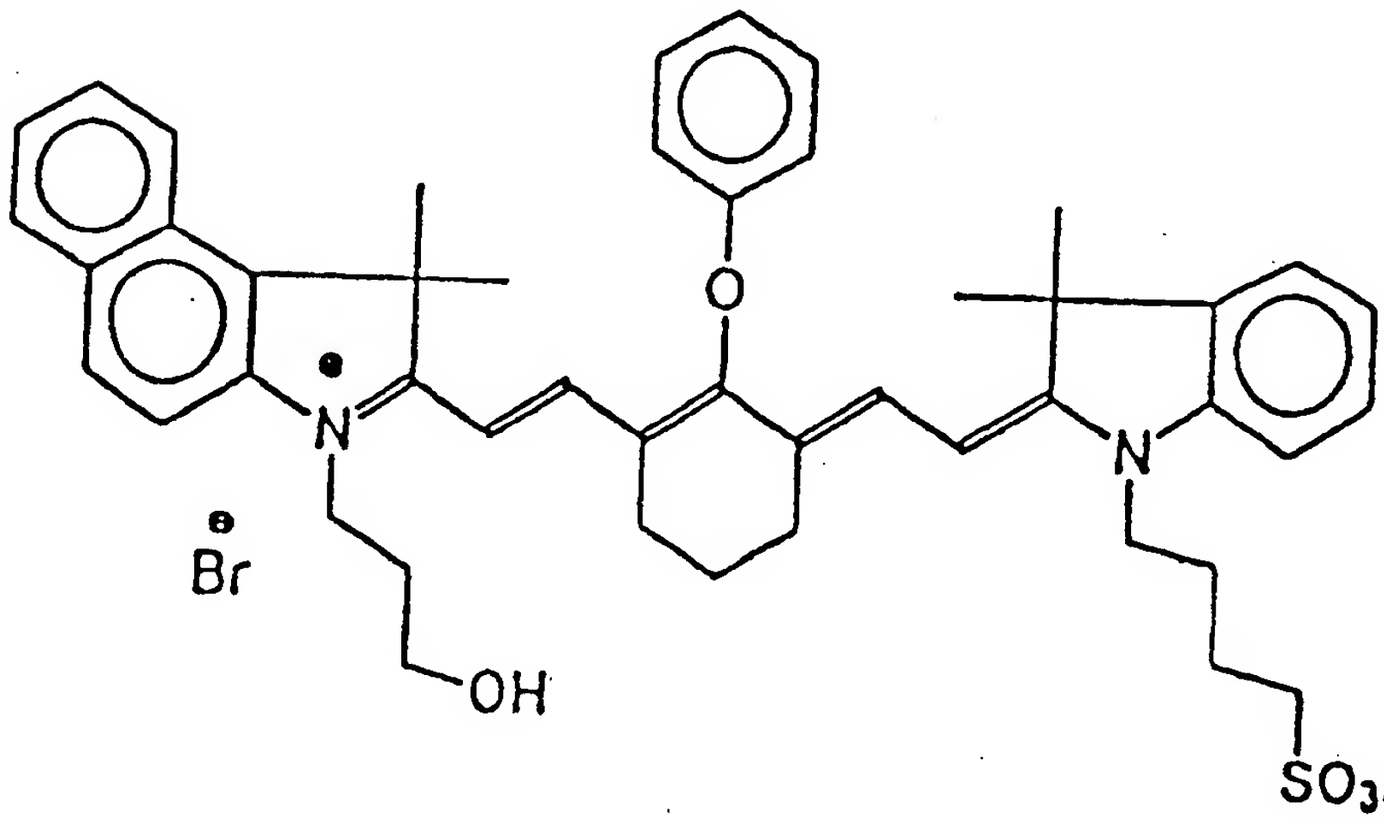
More specifically, the DNA samples marked with fluorescent dye having absorbance and fluorescense maxima at near infrared or infrared wavelengths when combined with the DNA are applied at a plurality of locations in a gel electrophoresis slab for electrophoresing in a plurality of channels through a gel electrophoresis slab. An electrical potential is established across said gel electrophoresis slab wherein DNA

samples are resolved in accordance with the size of DNA fragments in said gel electrophoresis slab into fluorescently marked DNA bands. The separated samples are scanned photoelectrically while they are in the slab with a laser diode and a sensor wherein the laser scans with near infrared or infrared scanning light at a scanning light frequency within the absorbance spectrum of said fluorescently marked DNA samples and light at the emission frequency of the marked DNA is sensed.

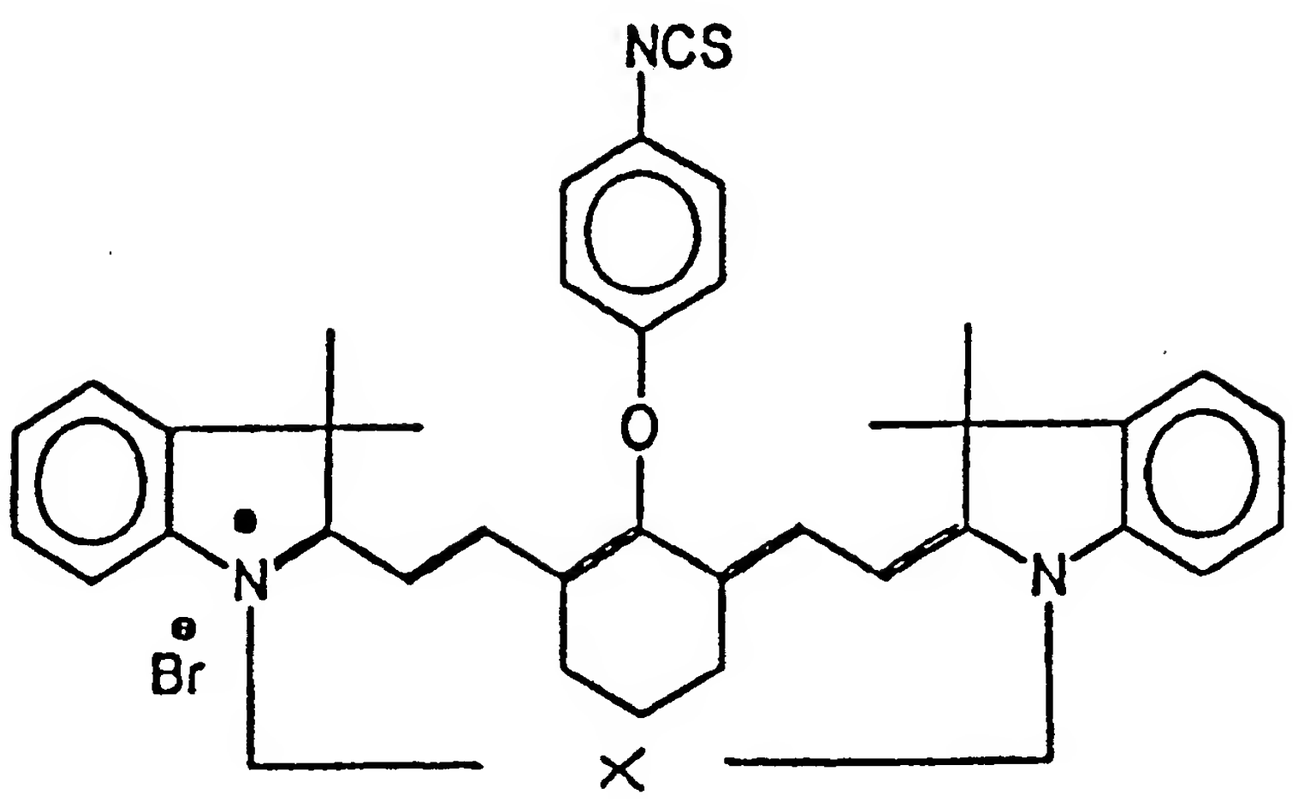
In this method, suitable dyes comprise:



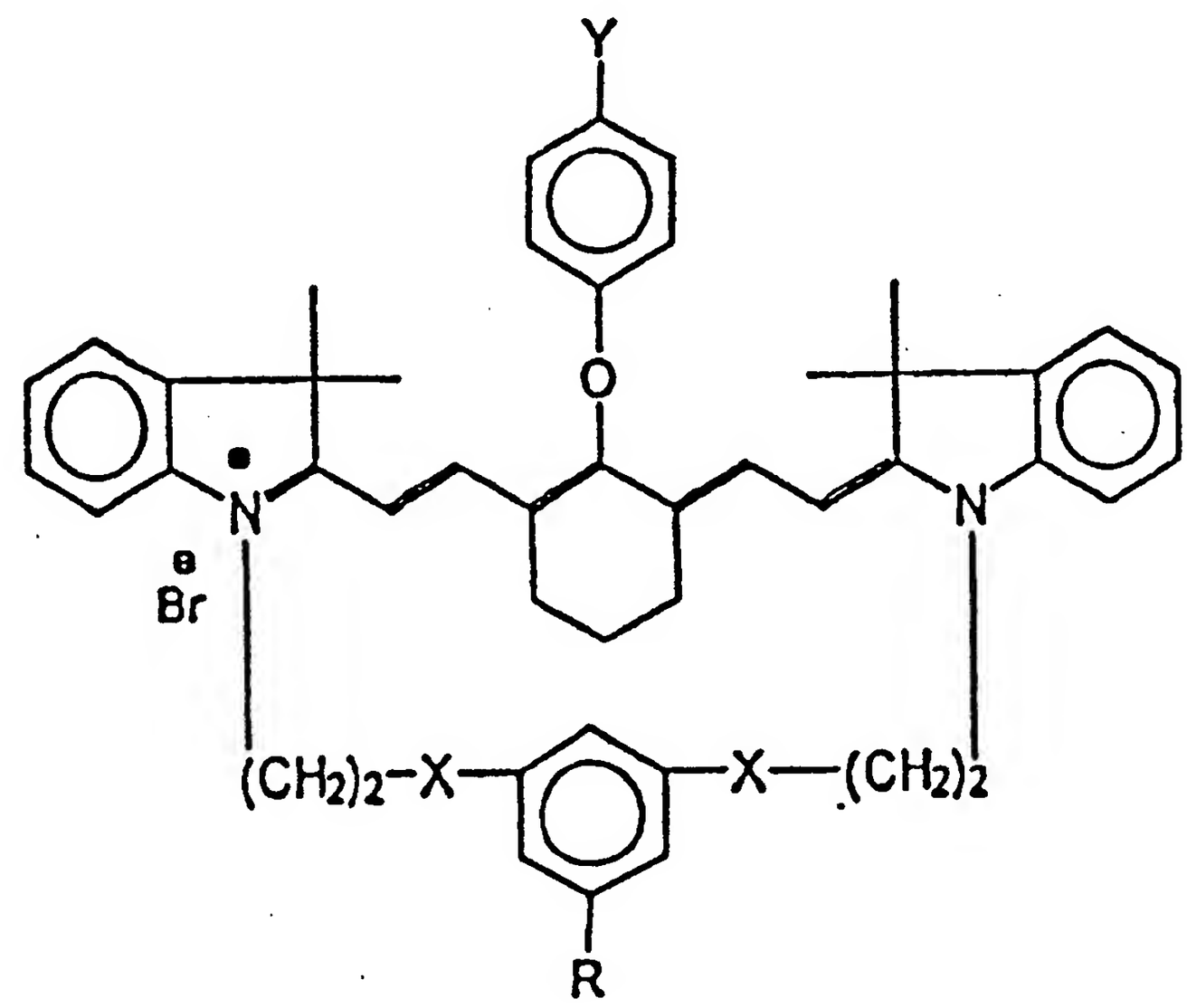
or



or



where X is $(CH_2)_n$; n = 4-10 or X is $-CH_2 - CH_2 - 0 - CH_2 - CH_2 - 0 - CH_2 - CH_2 -$;
or



wherein X is one of O or NH; Y is one of NCS or H; R is one of H, NCS, CH_2OH, CH_2NCH or COOH.

From the above summary, it can be understood that the techniques for the sequencing of fluorescence labeled DNA of this invention have several advantages, such as: (1) because the dyes have their emission spectra in the infrared or near infrared light spectrum; small inexpensive infrared diode lasers may be used; and (2) they are characterized by relatively low noise.

The above noted and other features of the invention will be better understood from the following detailed description when considered with reference to the accompanying drawings in which:

FIG. 1 is a perspective view of an embodiment of sequencer usable in the invention;

FIG. 2 is a sectional view taken through lines 2-2 of FIG. 1;

FIG. 3 is a sectional view of a portion of FIG. 1 taken through lines 3-3;
FIG. 4 is an exploded perspective view of a portion of the embodiment of FIG. 2;
FIG. 5 is an enlarged view, partly broken away, of a portion of the embodiment of FIG. 2; and
FIG. 6 is a block diagram of a circuit that may be used for coordination of a sensor, scanner drive and laser used.

DETAILED DESCRIPTION

The sequencing of infrared fluorescence labeled DNA and the detection of the DNA after irradiation by infrared light from a laser diode is accomplished using an infrared label prepared for this purpose and either directly attached to the DNA or attached to probes or primers that will be attached to the DNA. In this specification the word "infrared" will be used at times to include near infrared wavelengths (700-900 nm) infrared (630-700 nm) and far infrared (900-3000 nm). The strands of DNA are continuously electrophoresed and identified for any of several purposes, such as for example: (1) DNA sequencing; and (2) analysis of strands varying in length as prepared by such techniques as restriction enzyme cutting or polymerase chain reaction (PCR).

The strands are marked with fluorescent labels that have their maximum fluorescense and their maximum absorbance at wavelengths of light in the infrared and near infrared region. The strands are irradiated with light in the infrared or near infrared region from a laser diode and the light emitted from the fluorescent labels is detected and used to obtain information about the DNA strands. The detector includes a light sensor which is preferably an avalanche photodiode sensitive to the near infrared light emission of the marker. It may include a filtering system having a pass band suitable for passing selectively the optimum emission of the fluorescent marker to the light sensor.

To mark the DNA strand, a dye is prepared having the desired properties or a known dye is modified. In the preferred embodiment a novel dye having the preferred spectrum, high absorption and fluorescence properties, and at least one reactive group enabling coupling to biomolecules such as DNA, proteins, and antibodies is synthesized.

The dye is synthesized, modified from a known dye or selected to have an absorbance band and an emission band within a region encompassing the near infrared and infrared regions when attached to a probe, primer, oligonucleotide or other molecule. The dye should provide high quantum yield in a band selected to reduce background noise. The preferred dyes for many applications calling for the labelling biomolecules are cyanine dyes having a NCS group on the dye that may react with the amino goup of the biomolecule to form a thiourea linkage.

In the preferred embodiment, cyanine dyes are synthesized. The preferred dyes are heptamethine cyanines which absorb light having wavelengths in the region of 750 to 820 nm (nanometers) efficiently (maximum absorbance wavelength). This wavelength is suitable for reducing background fluorescence in DNA sequencing and corresponds to the radiation wavelength of the GaAlAs diode laser which is 780 nm. The GaAlAs diode is used for scanning the gel electrophoresis sandwich used for DNA seqencing. Formulas 1-17 are typical synthesized or proposed dyes and formula 18 is a suitable modified dye.

Formula 1 shows synthesized cyanine dyes having NCS as a reactive group for attachment of a biomolecule. In this embodiment, when X is H, the maximum absorbance wavelength is between 787 and 801 nm, depending on the solvent and the maximum emission wavelength (maximum quantum yield and lambda flourescent light) is between 807 and 818 depending on the solvent. When X is $-OCH_3$, the maximum absorbance wavelength is 786 nm and the maximum emission wavelength is 806 nm. The quantum yield is high.

Formula 2 shows synthesized cyanine dyes some of which have a high quantum yield. For example, when R is ethyl or sulfonatobutyl and N is either 1 or 2, the maximum absorbance wavelength is between 762 and 778 nm. Depending on the solvent the maximum emission wavelength is between 782 and 800 nm.

In the synthesized dyes represented by formula 3, when R is SO_3Na, the maximum absorbance wavelength is between 773 and 778 nm and the maximum emission wavelength is between 789 and 806 nm.

In the molecule shown by formula 4, a phenoxy group permits an increase in excitation state of a larger number of electrons. This dye includes a functional hydroxy group to permit coupling to proteins such as DNA strands for the purposes of sequencing DNA.

As shown by formula 5, a still further increase in quantum yield is obtained by coupling the nitrogen groups on the two sides of the heterocyclic base by a polymethylene group including a number of methylene monomers chosen to maintain the distance between them. Generally, the methylene chain is between four and ten groups long, with nine being preferable. Similarly, ethers may be used with the same

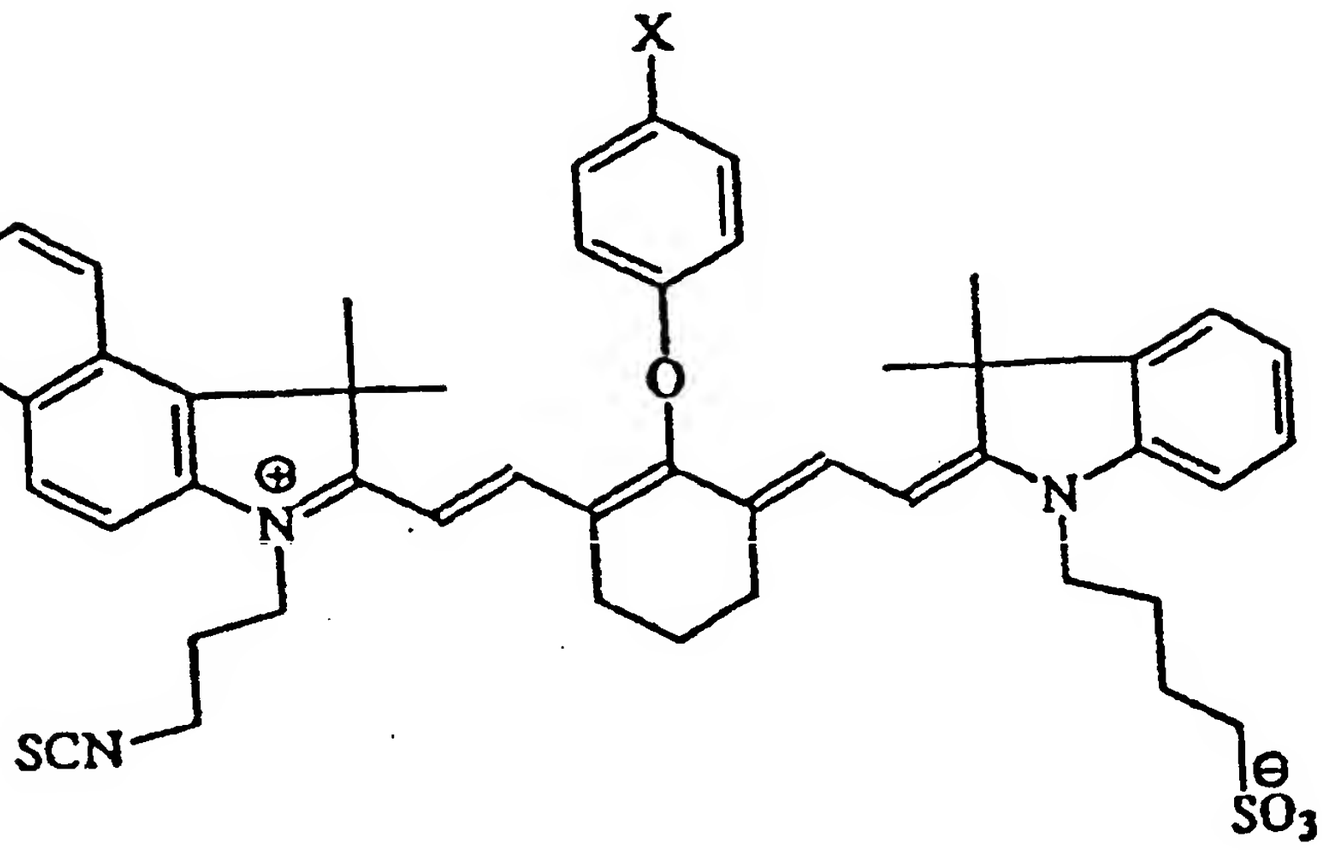
number of groups. The ethers are more stable. A functional group may be attached to provide convenient coupling. This arrangement has the proper spectrum and higher quantum yield.

A proposed series of molecules represented by formula 6 is expected to have increased solubility to permit its use in an aqueous medium through the use of an amino group or ether group attached to the benzene ring. In this series, X may be oxygen or NH, Y may be NCS (N-chlorosuccinimide) or H, and R may be H, NCS, CS, CH_2OH, CH_2NCS or COOH according to tables 1 or 2.

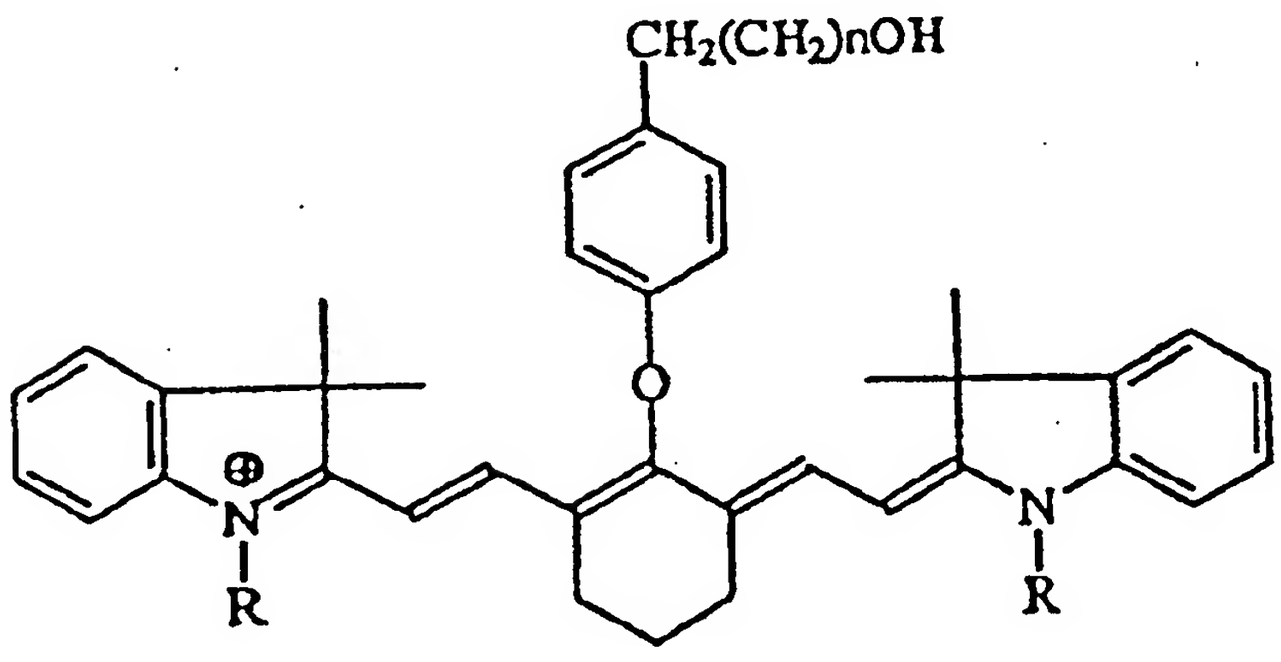
In formula 7: R may be ethyl, 4-sulfonatobutyl, 3=amino propyl phthalimidobutyl, hexyl and pentyl carboxylates; Z may be S, O, CMe_2, Y may be H, SO_3Na, SO_3ET_3NH, OCH_3, NO_2 CO_2Na or CO_2ET_3NH and X may be H, $N(CH_2CO_2CH_3)_2$, CH_2CH_2OH, CHS, $CH_2CH_2CH_2OH$ or -(CH_2)n-OH where the number n= any number between and including 8-12.

In the series of proposed molecules represented by formula 8:

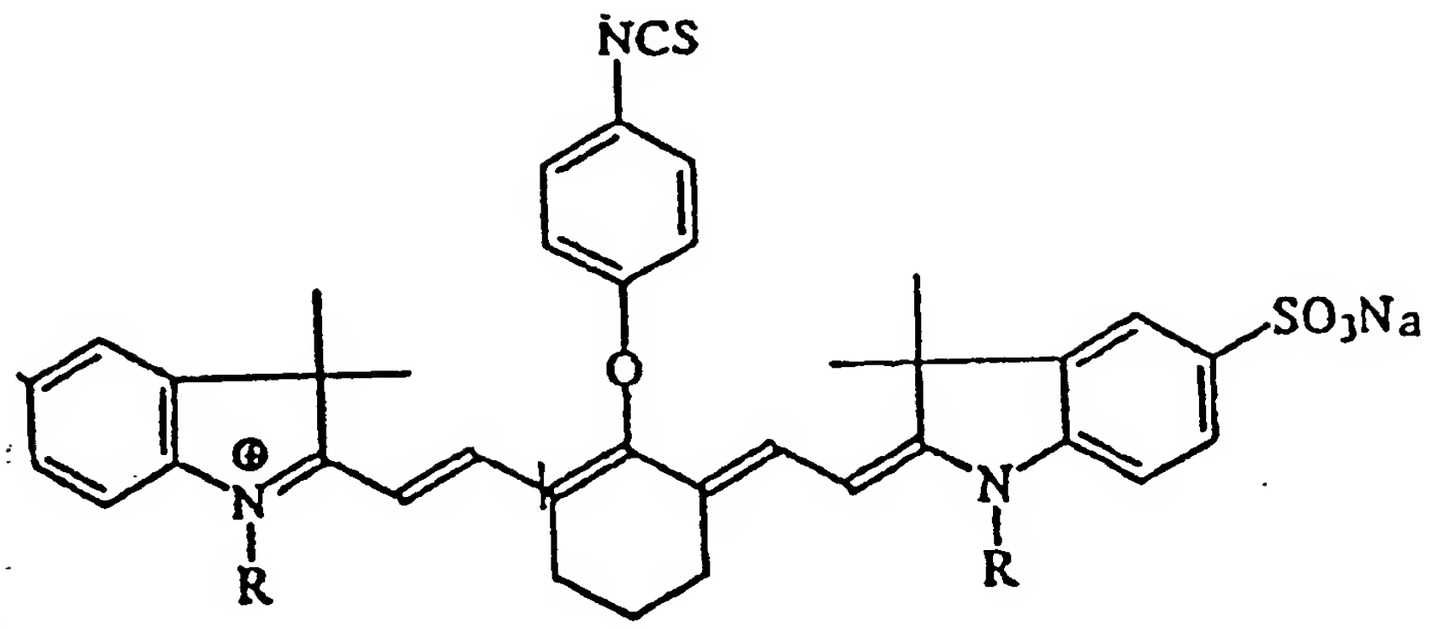
FORMULA 1



FORMULA 2



FORMULA 3



FORMULA 4

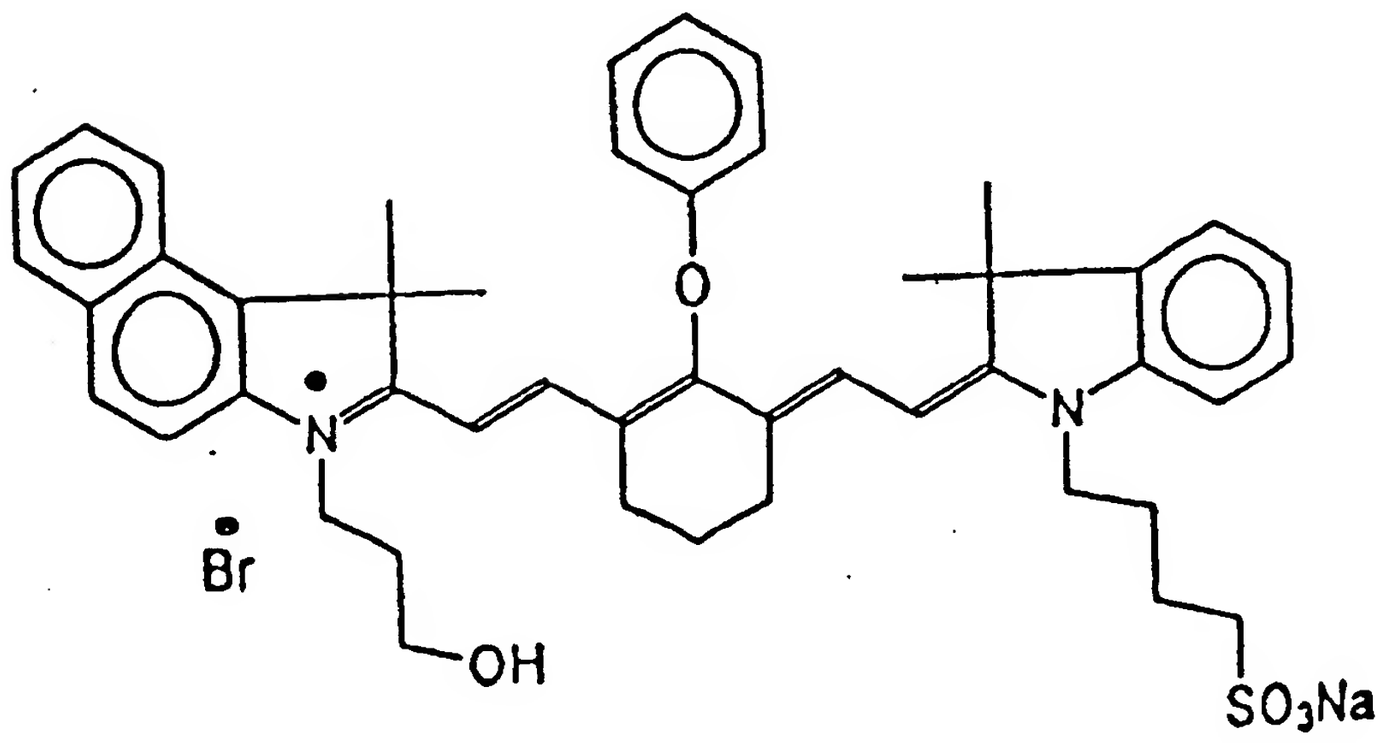


TABLE 1

Dye formula 4 series, when X = 0		
X	Y	R
0	NCS	H
0	H	NCS
0	H	CH_2OH
0	H	CH_2NCS
0	H	COOH

TABLE 2

Dye formula 4 series, when X = NH		
X	Y	R
NH	NCS	H
NH	H	NCS
NH	H	CH_2OH
NH	H	CH_2NCS
NH	H	COOH

(1) When X and R are hydrogen, the absorbance wavelength is 767 nanometers and the emmission wavelength is 787 nanometers, a high quantum yield.
(2) When X is equal to MeO and R is hydrogen, the absorbance wavelengths are slightly higher such as between 767 and 769 nanometers and the emission wavelength is slightly greater but the yield is slightly reduced.
(3) When X is a nitro group and R is hydrogen, the absorbance wavelength increases as well as the emission wavelength, with the absorbance wavelength being between 770 and 774 nanometers and emission wavelength being between 792 and 798 wavelengths with a high quantum yield.
(4) When X is NCS and R is hydrogen, the absorbance and emission wavelengths drop to between 768 and 769 nanometers for the absorbance wavelength and 788 nanometers for the emission wavelength with a quantum yield that is high.
(5) When X is NCS and R is OM the absorbance maximum wavelength is between 790 and 796 nm and the maximum emission wavelength is 814 nm.
(6) When X is C_3H_6OH and R is hydrogen, the maximum absorbance wavelength is between 766 and 768 nm and the maximum emission wavelength is between 788 and 790 nm.
(7) When X is NCS and R is SO_5 Na, the maximum absorbance wavelength is between 773 and 778 nm and the maximum emission wavelength is between 789 and 806 nm.

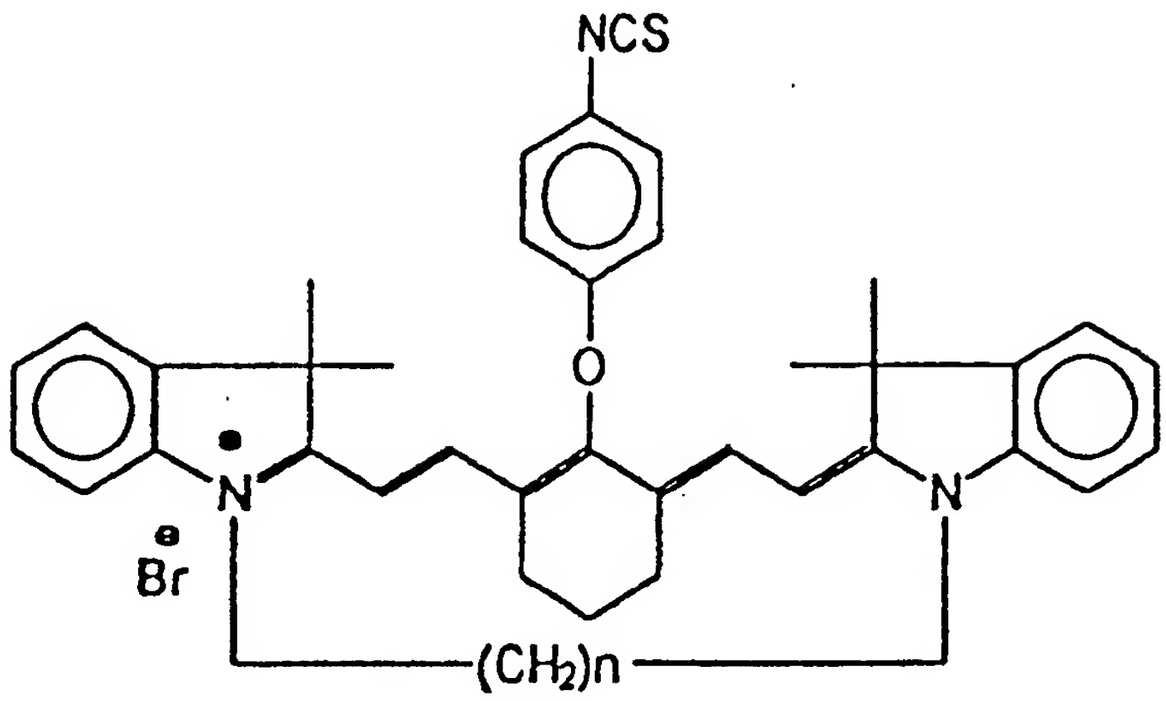
In the series of molecules of formula 9:

(1) When R is OMe, and X is NCS, the absorbance wavelength is 790 to 796 nanometers and the emission wavelength is 814 nanometers. The quantum yield is good.
(2) When R is H, and X is -$CH_2CH_2CH_2$-OH (hydropropyl) the aborbance wavelength is slightly reduced to between 766 and 768 nanometers and the emission wavelength is reduced to between 788 and 790 nanometers but the quantum yield is increased.
(3) When R is SO_3Na and X is NCS, the absorbance wavelength is reduced to 773 to 778 nanometers and the emission wavelength is reduced slightly to 789 to 806 nanometers. It produces a high quantum yield.

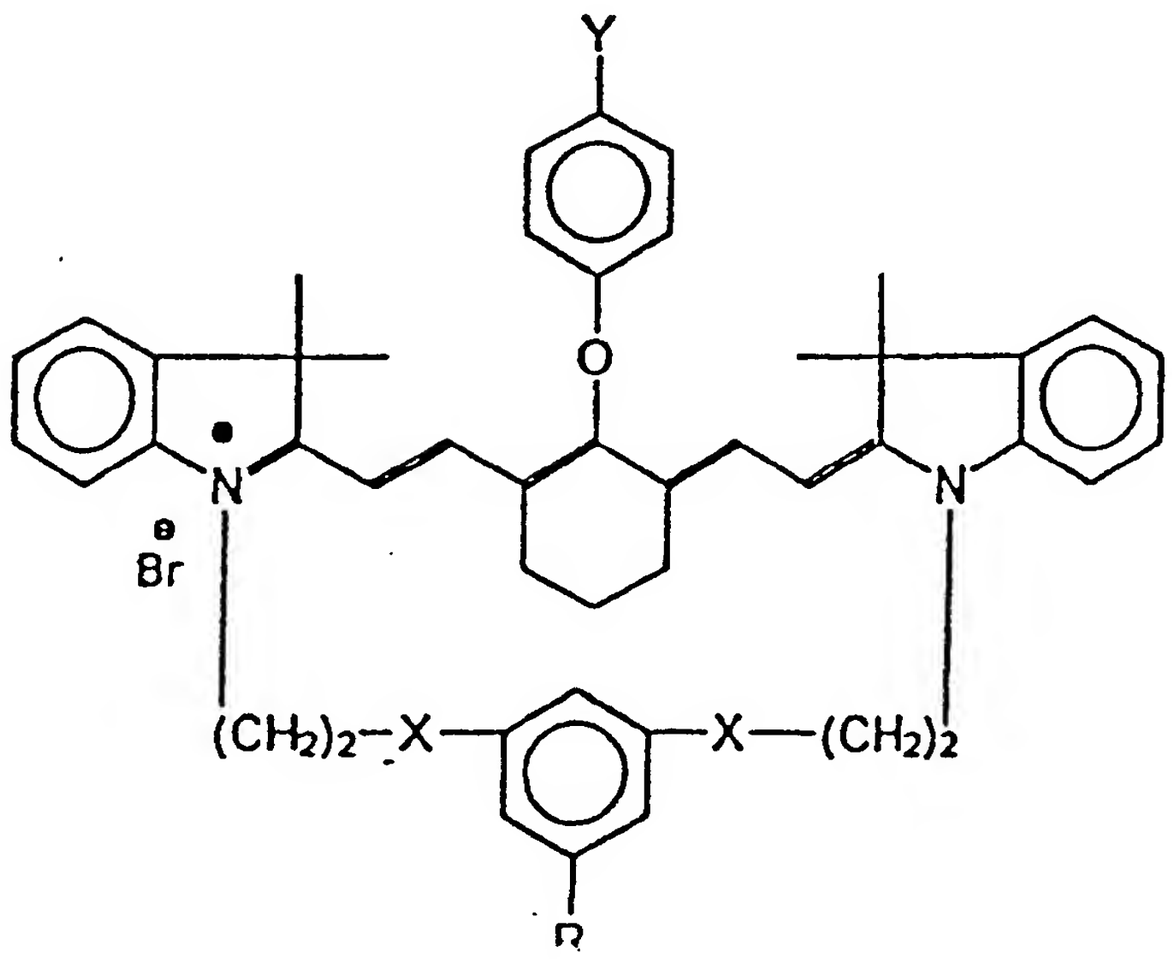
In the series of molecules of formula 10:

(1) When X is NCS the absorbance wavelength is between 764 and 770 nanometers and the emission

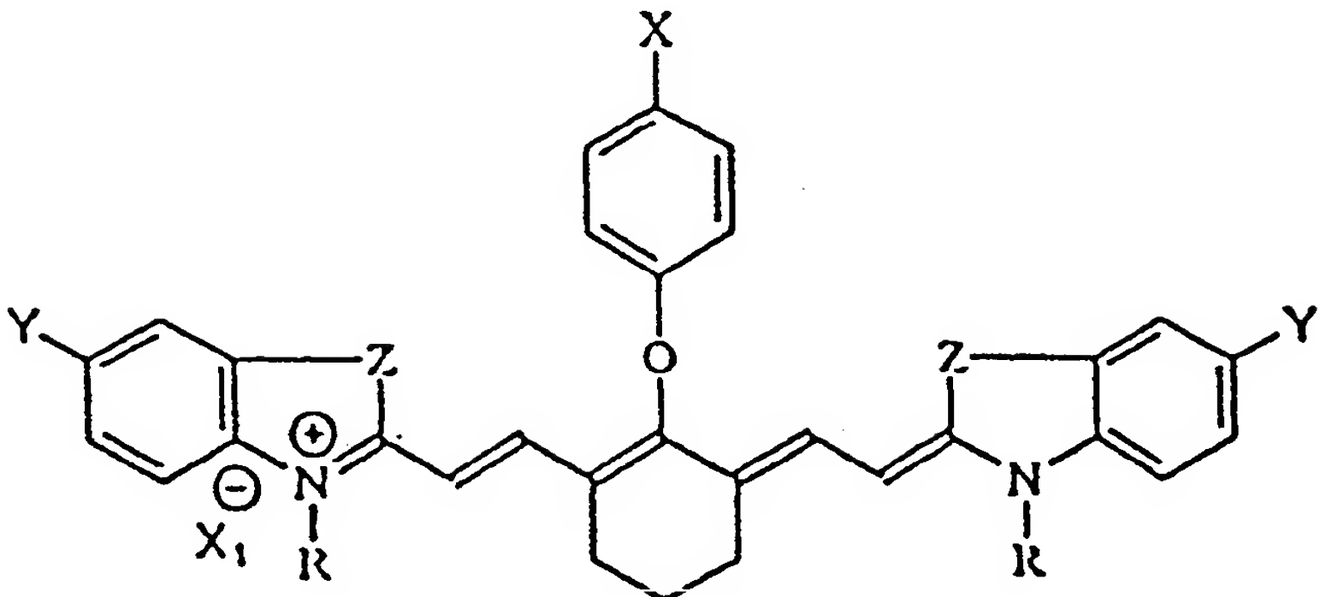
FORMULA 5



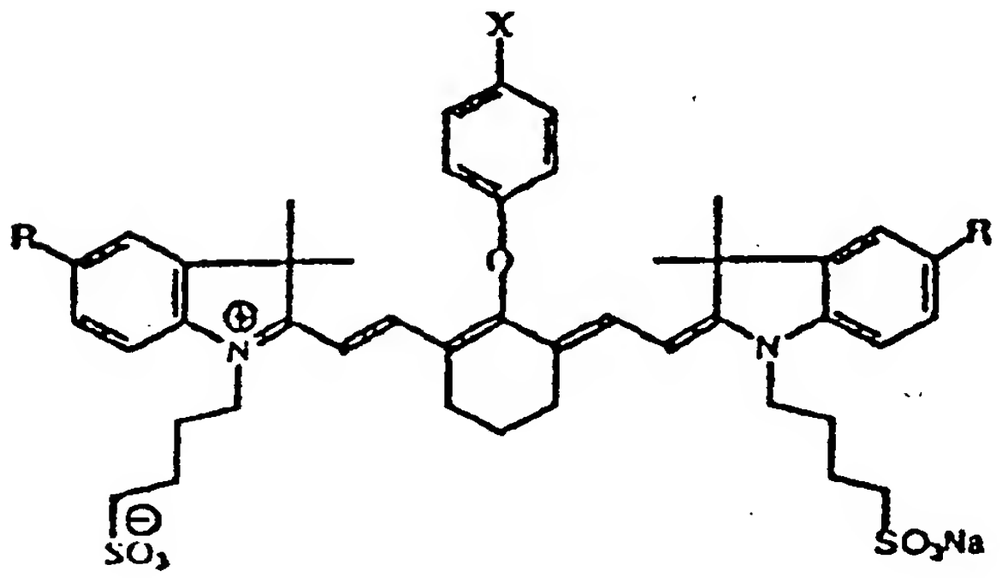
FORMULA 6



FORMULA 7



FORMULA 8



wavelength is between 786 and 802 nanometers with a moderate quantum yield.

(2) When X is $CH_2CH_2CH_2OH$ the absorbance wavelength is slightly reduced to between 762 and 770 nanometers and the emission wavelength is between 782 to 790 nanometers with a high quantum yield increase.

The series of molecules represented by formula 11 have been synthesized and are effective as IR labels when R1, R2 and X are as shown in table 3.

The series of molecules represented by formula 12 have been synthesized and are effective as IR labels when R_1 and X are as shown in table 4.

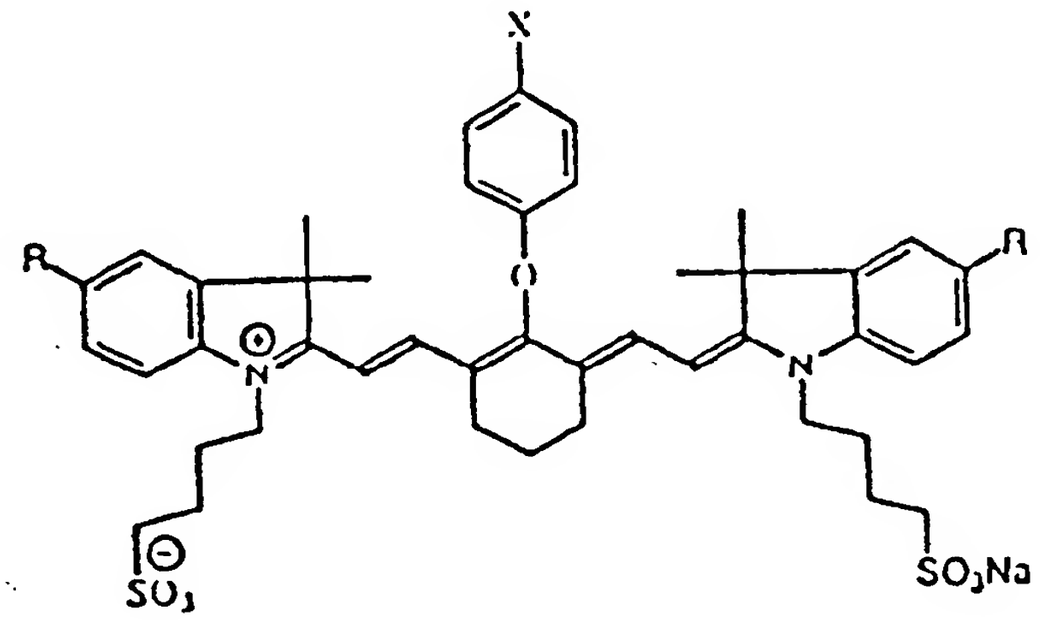
The series of molecules represented by formula 13 have been synthesized and are effective as IR labels when R_1, R_2 and X are as shown in table 5.

The series of molecules represented by formula 14 have been synthesized and are effective as IR labels when R_1, R_3 and R_2 are as shown in table 6.

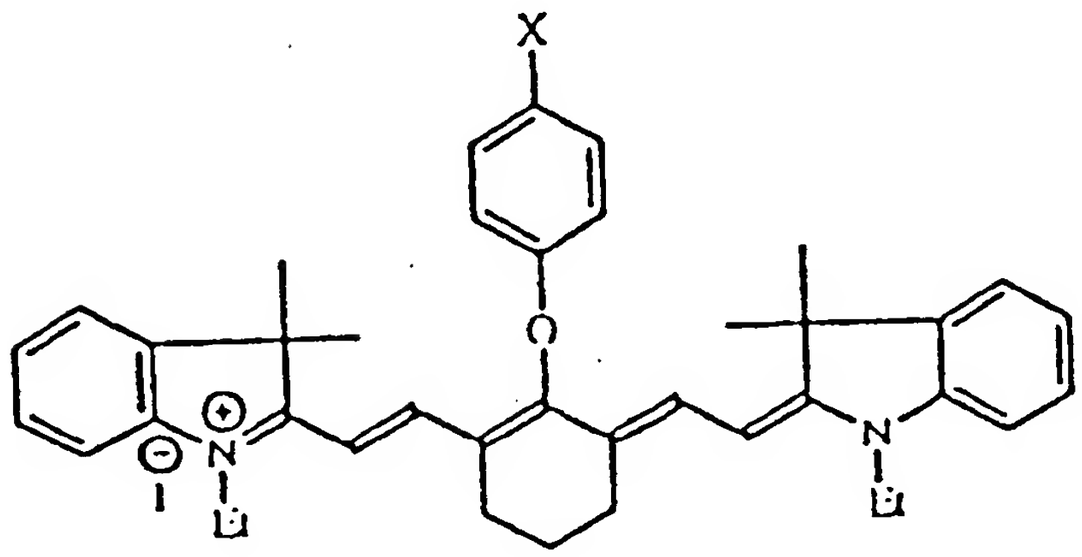
The series of molecules represented by formula 15 have been synthesized and are effective as IR labels when R and X are as shown in table 7.

The series of molecules represented by formula 17 have been synthesized and are effective as IR labels when R is shown in table 8.

FORMULA 9



FORMULA 10



A series of dyes is obtained by modifying a commercial dye, shown as formula 18, with R being - CH_2-CH_3. This series is close to having the desired wavelength of maximum fluorescence. The wavelength of

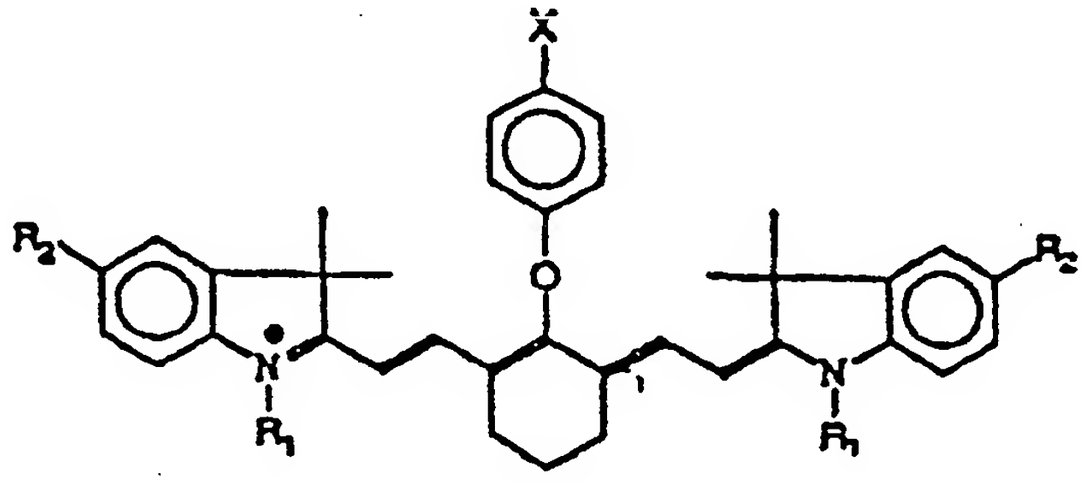
maximum absorbance may be modified by changing the functional group R. The unmodified dye may be obtained from Laboratory and Research Products Division, Eastman Kodak Company, Rochester, New York, 14650. It is advertised in Kodak publication JJ-169.

In the preferred embodiment, the fluorescence maximum wavelength is about 819 nanometers and the detector is adjusted to receive this wavelength and not others by appropriate filtering. The absorption maximum is selected to be different and to correspond to the preferred available laser diode emission. For example, in this formula, R may be any of the following four groups, depending on the desired wavelength of the absorbed light, which are:

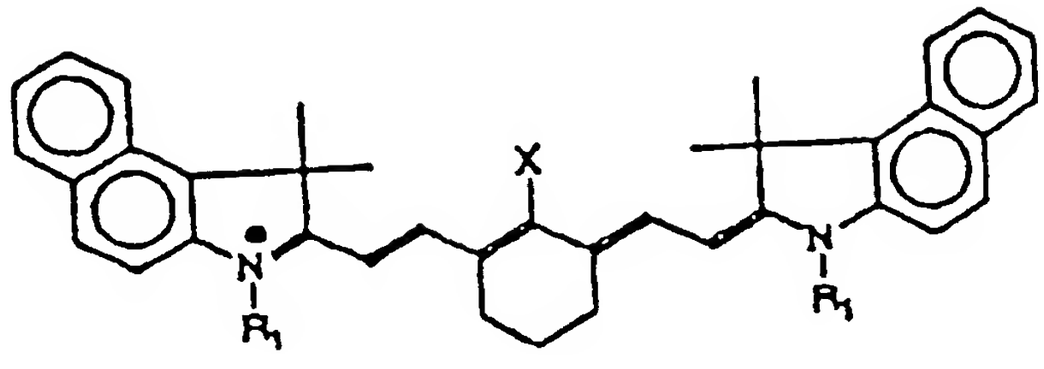
(1) $-CH_2-CH_2-OH$ for an excitation wavelength of 796 nanometers;

(2) $-CH_2-CH_2--CH_2-OH$ for an excitation wavelength of 780 nanometers, which is the preferred embodiment;

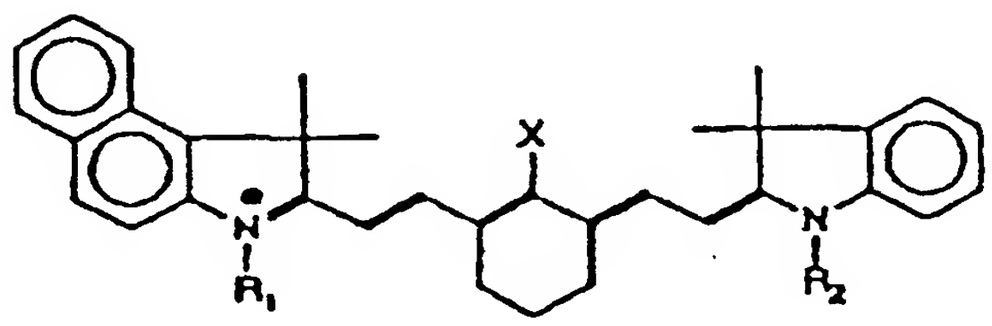
FORMULA 11



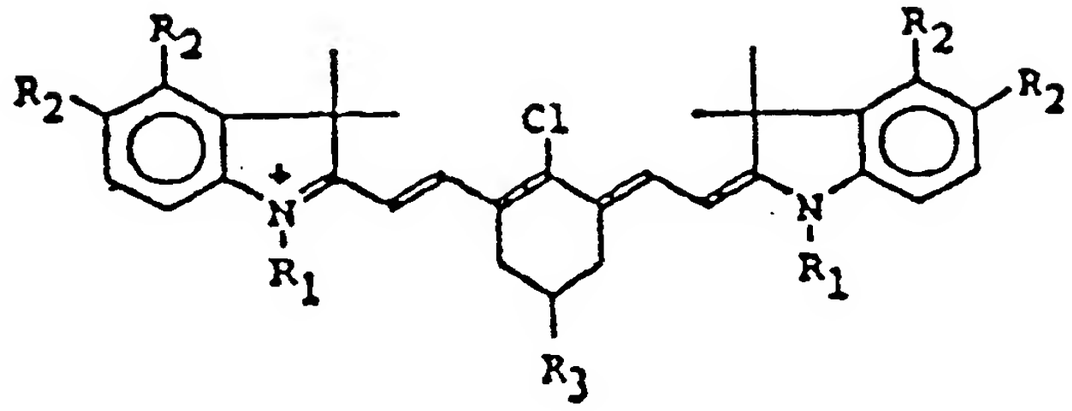
FORMULA 12



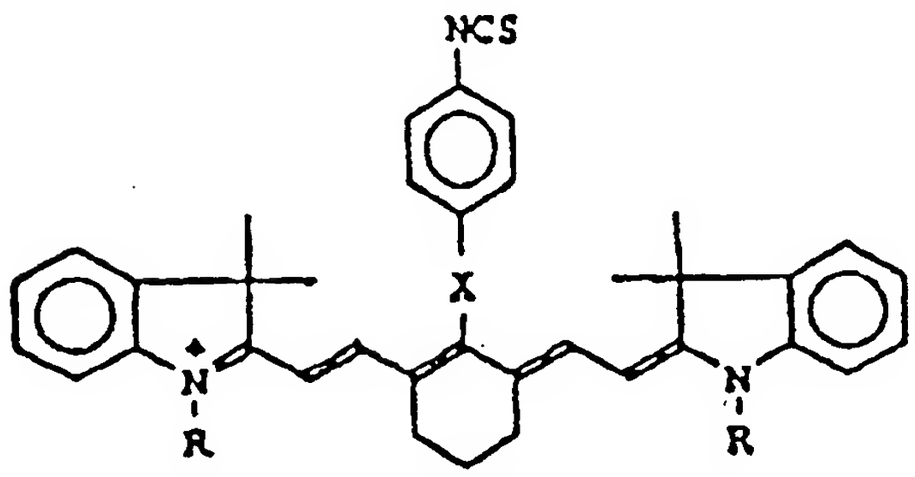
FORMULA 13



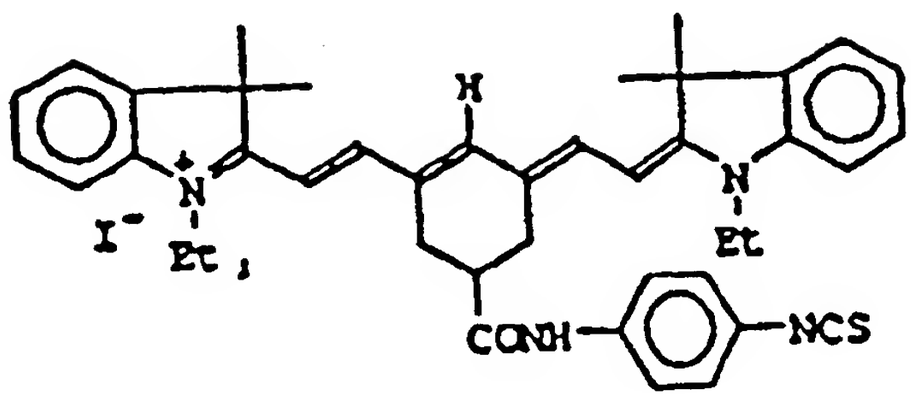
FORMULA 14



FORMULA 15



FORMULA 16



FORMULA 17

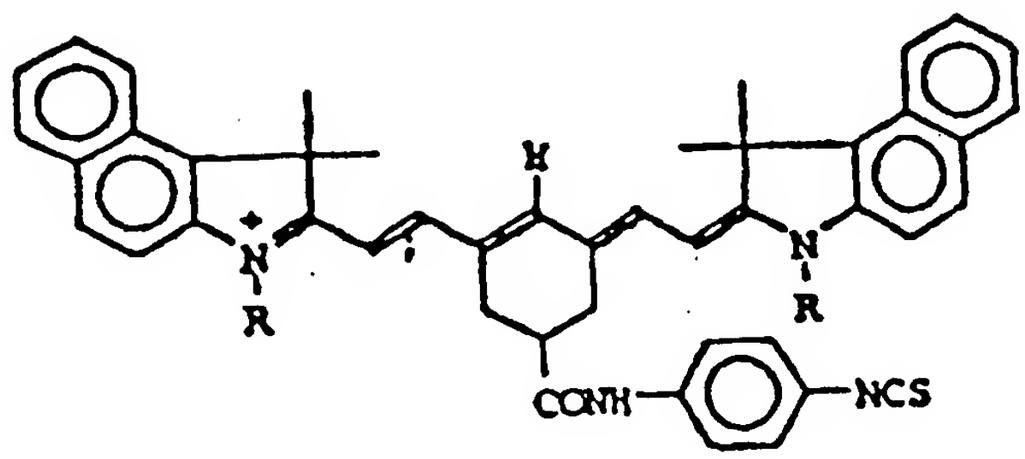


TABLE 3

R1	R2	X
$-C_4H_8SO_3(Na)$	H	NCS
$-C_4H_8SO_3(Na)$	H	OCH_3
$-C_4H_8SO_3(Na)$	H	NO_2
$-C_4H_8SO_3(Na)$	H	H
$-C_2H_5(I)$	H	OCH_3
$-C_2H_5(I)$	OCH_3	OCH_3
$-C_4H_8SO_3(Na)$	OCH_3	OCH_3
$-C_4H_8SO_3(Na)$	OCH_3	NCS
$-C_3H_8NCS(Br)$	H	H
C_2H_5I	H	NCS
$C_4H_8SO_3Na$	SO_3Na	NCS
$C_2H_5(I)$	SO_3Na	NCS

Table 4

R_1	X
$-C_2H_5(I)$	OPhNCS
$-C_4H_8SO_3(Na)$	OPhNCS
$-C_3H_8NCS(Br)$	OPh
phthalimide butyl	OPhNCS

Table 5

R_1	R_2	X
$-C_3H_6NCS(Br)$	$-C_4H_8SO_3H$	Cl
$-C_3H_6NCS(Br)$	$-C_2H_5$	Cl
$-C_3H_6NCS$	$-C_4H_8SO_3$	OPh
$-C_3H_6NCS(Br)$	$-C_2H_5$	OPh
C_3H_6NCS	$C_4H_8SO_3$	OPh(pMeO)

Table 6

R_1	R_3	R_2
Et(I);	H	H
$(CH_2)_4SO_3$ (Na);	H	H
Et(I);	$COOCH_3$	H
Et(I)	$COOCH_3$	$-(CH)_4-$
$(CH_2)_4SO_3$ (Na)	$COOCH_3$	-(CH)4-

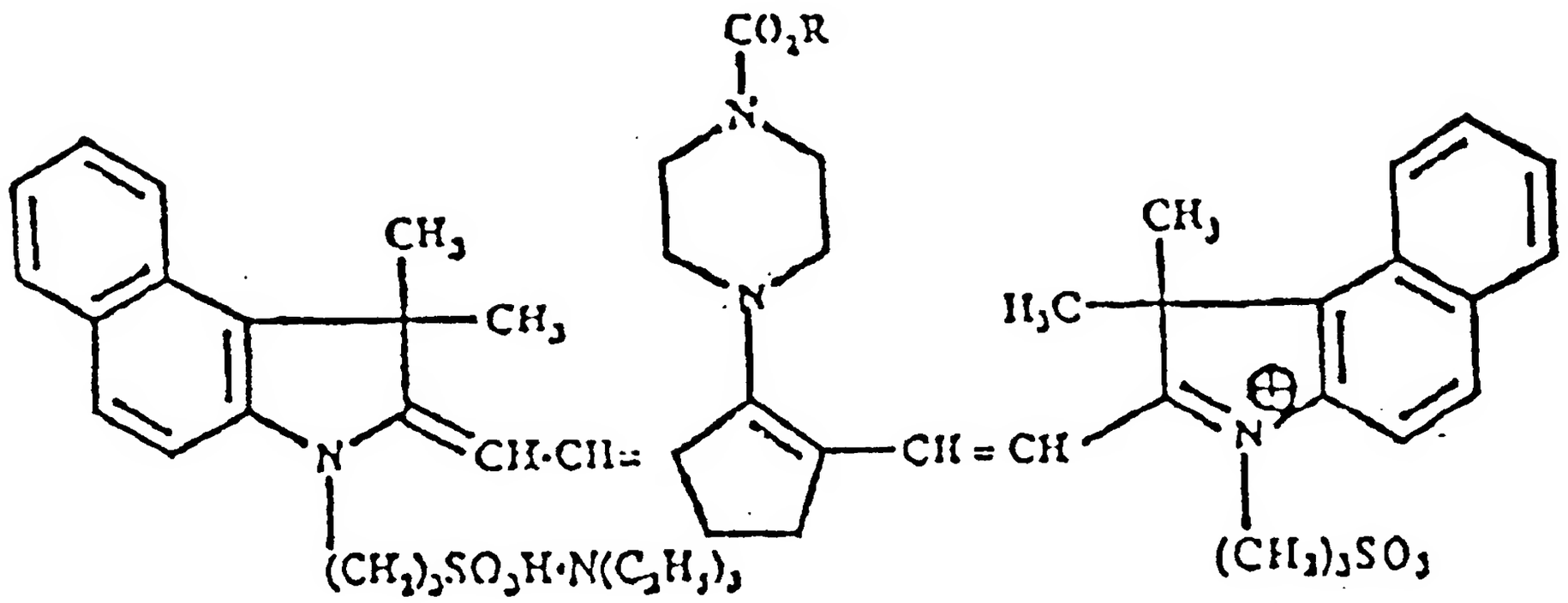
TABLE 7

R	X
$Et(I^-)$	O
$(CH_2)_4SO_3^-(Na^+)$	O
$Et(I^-)$	S
$(CH_2)_4SO_3^-(Na^+)$	S

TABLE 8

R
$Et(I^-)$
$(CH_2)_4SO_3^-(Na^+)$

FORMULA 18



(3) $-CH_2-CH_2--CH_2-CH_2-CH_2-CH_2-OH$ for an excitation wavelength of 745 nanometers; and

(4) $-CH_2-CH_2-O-CH_2-CH_2-O-CH_2-O-CH_2-OH$ for an excitation wavelength of 790 nanometers.

In another embodiment, the infrared dyes are selected to provide at least four dyes with each dye having its exitation and/or radiation spectra spaced sufficiently from each other dye so the flourescence from each dye can be electronically distinguished from each other dye either by the wavelength that excites it into fluorescence or the wavelength at which it fluoresces or both. The spacing is maintained sufficiently close to be excited by laser diodes. The functional groups may be modified as explained above for spacing between the absorption maxima and fluorescence maximum. The dyes may be incorporated in probes and primers for attachment to oligonucleotides as described in Ruth, Jerry L. (1984) DNA 3, 123. The -OH group provides appropriate linkage.

There are many dyes suitable for such modification such as for example: (1) 3,3'-Diethylthiadicarbocyanine Iodide; (2) 3,3'-Diethylthiatricarbocyanine Perchlorate; (3) 3,3' Diethyloxatricarbocyanine Iodide; (4) 1,1',3,3,3'-Hexamethylindotricarbocyanine Perchlorate; (5) 1,1'-Diethyl-2,2'-dicarbocyanine Iodide; (6) 3,3'-Diethylthiadicarbocyanine Iodide; (7) 3,3'-Diethyloxatricarbocyanine Iodide; (8) 1,1',3,3,3',3'-Hexamethylindotricarbocyanine Perchlorate; (9) 1,1',3,3,3',3'-Hexamethylindotricarbocyanine Iodide; and (10) Indocyanine Green.

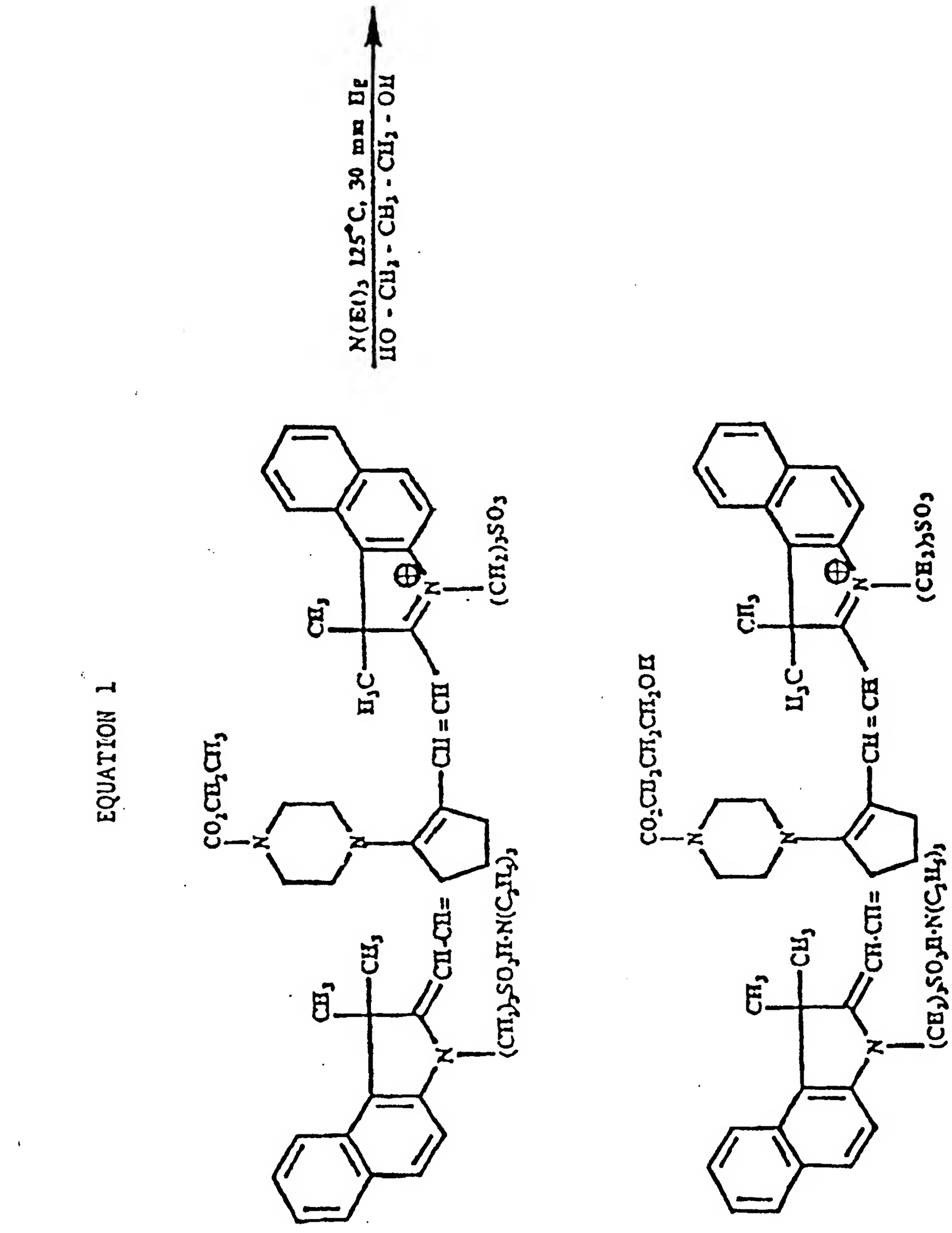
In the one embodiment, the dye has the formula shown in formula 18, with R being $-CH_2-CH_3$. This dye is close to having the desired wavelength of maximum fluorescence and the wavelength of maximum absorbance may be modified by changing the functional group R. The unmodified dye may be obtained from Laboratory and Research Products Division, Eastman Kodak Company, Rochester, New York 14650. It is advertised in the Kodak laser dyes, Kodak publication JJ-169.

The modifications with 1,3-propanediol can be made in a manner known in the art as illustrated by equation 1. For example, changes occur when different esters are formed replacing the ethyl alcohol in the original dye molecule (R equal $-CH_2-CH_3$ of formula 18). If different glycol esters are formed, absorption maxima of these new near infrared dyes shift to the longer wavelengths. Moreover, new dyes may be synthesized rather than modifying existing dyes in a manner known in the art.

The absorption maximum is dependent on the distance of the O atoms in the glycol functional group. However, the fluorescence maxima of these new near infrared dyes are practically at same wavelength of the dye of formula 18, i.e. 819 nm. This indicates that only the excitation process has changed, i.e. to what energy level the transition occurs. The lowest vibronic level of first excited state remains unchanged. The absorption maxima of several such esters are: (1) ethylene glycol 796 nm (nanometers); (2) 1,3-Propanediol 780 nm; (3) 1,4-Butanediol 754 nm; (4) 1,6-Hexanediol 744 nm; (5) Triethylene glycol (#4) 790 nm; and (6) IR-144 ($R = CH_2-CH_3$) 742 nm.

In the preferred embodiment, the fluorescence maximum wavelength is about 819 nanometers and the detector is adjusted to receive this wavelength and not others by appropriate filtering. The absorption maxima is selected to be different and to correspond to the preferred available laser diode emission. For example, in formula 18, R may be any of the

EQUATION 1



following four groups, depending on the desired wavelength of the emitted light, which are:

(1) $-CH_2-CH_2-OH$ for an emission wavelength of 796 nanometers;

(2) $-CH_2-CH_2--CH_2-OH$ for an emission wavelength of 780 nanometers, which is the preferred embodiment;

(3) $-CH_2-CH_2-CH_2-CH_2-CH_2-CH_2-OH$ for an emission wavelength of 745 nanometers;

(4) -CH_2-CH_2-O-CH_2-CH_2-O-CH_2-O-CH_2-OH for an emission wavelength of 790 nanometers; and
(5) -CH_2-CH_2-SH for an emission wavelength of 810 nanometers.

The synthesis of these dyes is illustrated by chart 1, which shows the synthesis of formulas 7, 8, 9, 10, 11. In the chloro intermediate in chart 1, R_1 and R_2 may have the values shown in table 9. Chart 2 shows the synthesis of formula 12. Chart 3 shows the synthesis of formulas 1 and 13. Chart 5 shows the synthesis of formula 14. Chart 6 shows the synthesis of formula 15. Chart 7 shows the synthesis of formula 16. Chart 8 shows the synthesis of formula 17. In these charts X is a halogen.

Table 9

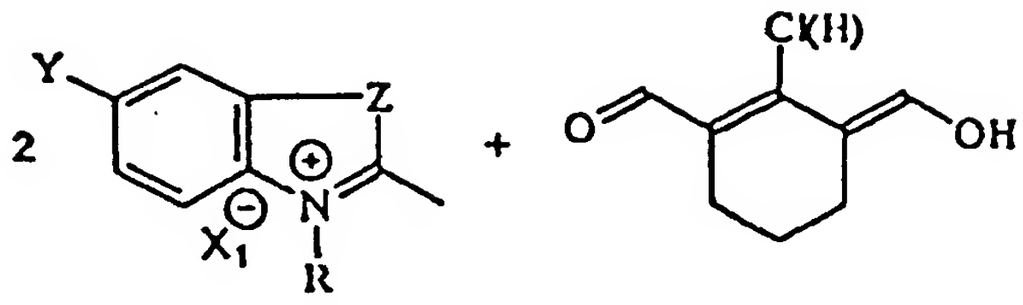
R	Y
$C_4H_8SO_3(H)$	H
C_2H_5	H
$C_3H_6NH_2HBr$	H
$C_4H_8SO_3(H)$	SO_3NC
$C_4H_8SO_3(H)$	OCH_3
C_2H_5	OCH_3

In still another embodiment, the infrared dyes are selected to provide at least four dyes with each dye having its excitation and/or radiation spectra spaced sufficiently from each other dye so the flourescence from each dye can be electronically distinguished from each other dye either by the wavelength that excites it into fluorescence or the wavelength at which it fluoresces or both. The spacing is maintained sufficiently close to be excited by laser diodes. The functional groups may be modified as explained above for spacing between the absorption maxima and fluorescence maximum.

In each of the embodiments, the dyes may be incorporated in probes and primers for attachment to oligonucleotides as described in Ruth, Jerry L. (1984) DNA 3, 123. The -OH group provides appropriate linkage to conventional probes by reaction with the appropriate reactive group such as primary amine, carboxylic acid groups and the like but near infrared dyes can be modified to have reactive groups other than the -OH for this purpose.

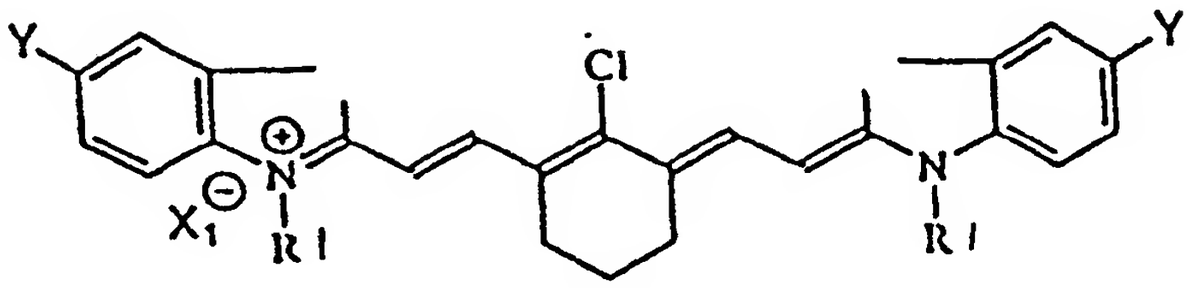
In FIG. 1, there is shown a perspective view of an embodiment of sequencer in which the method of this invention may be performed. This sequencer is described in structure and operation in the

CHART 1



$-H_2O$
No catalyst
n-butanol/benzene
(7:3)

Chloro Intermediate



4-substituted Phenol/DMF
NaH/0-5°C.

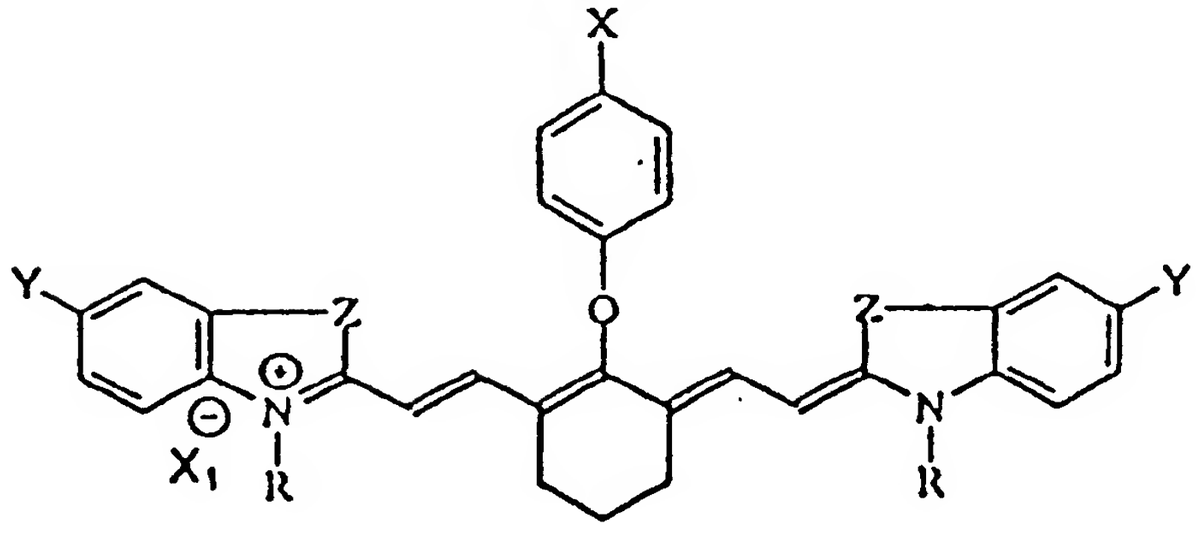


CHART 2

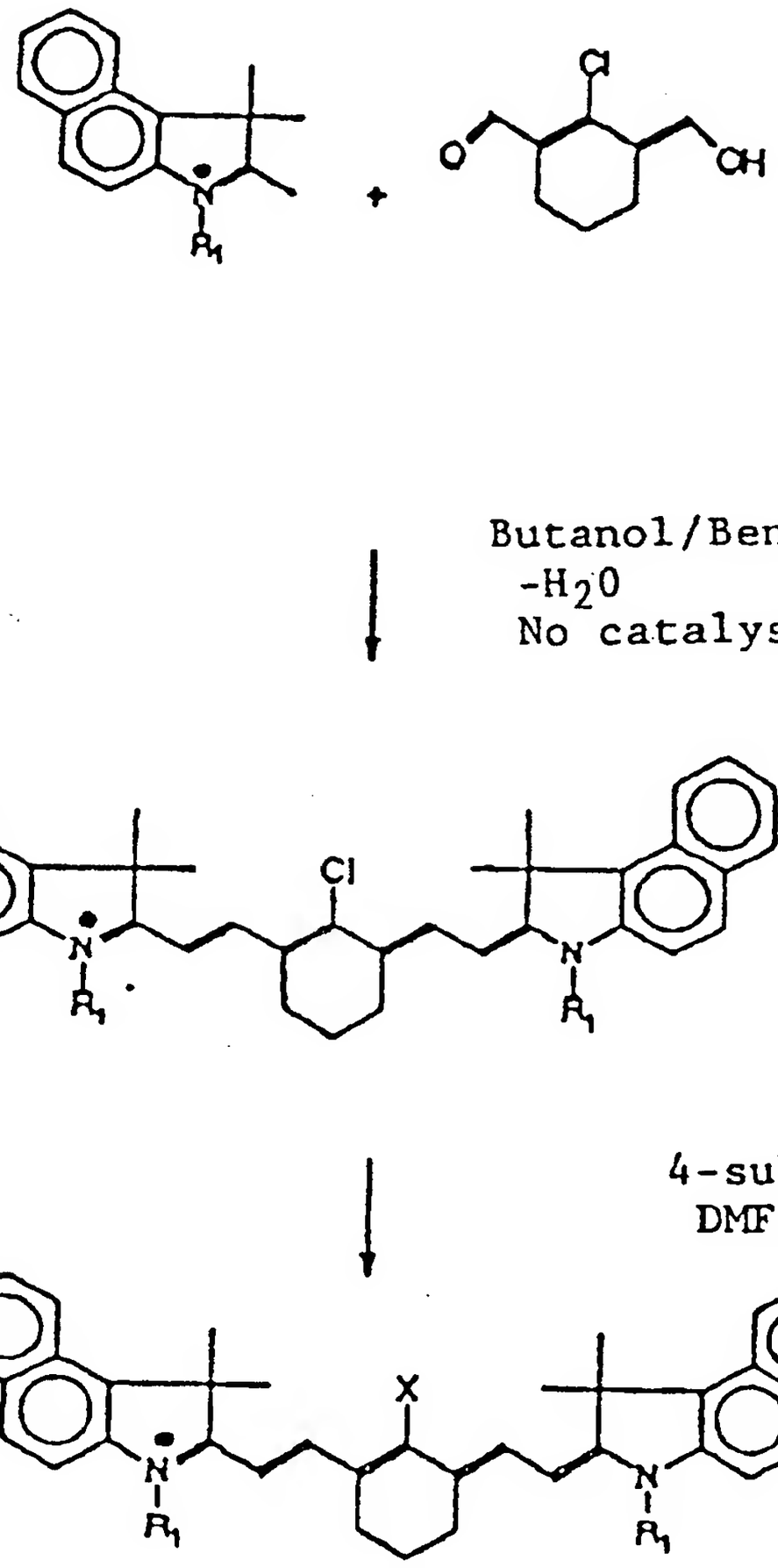
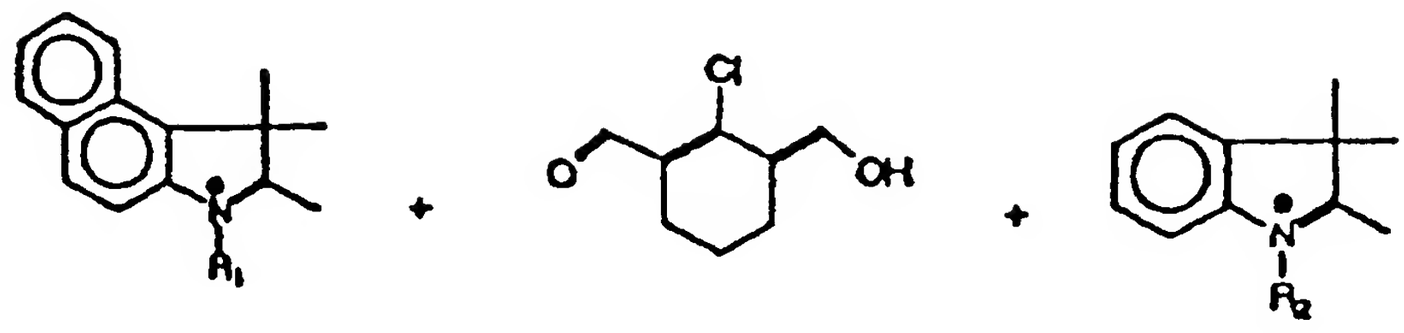
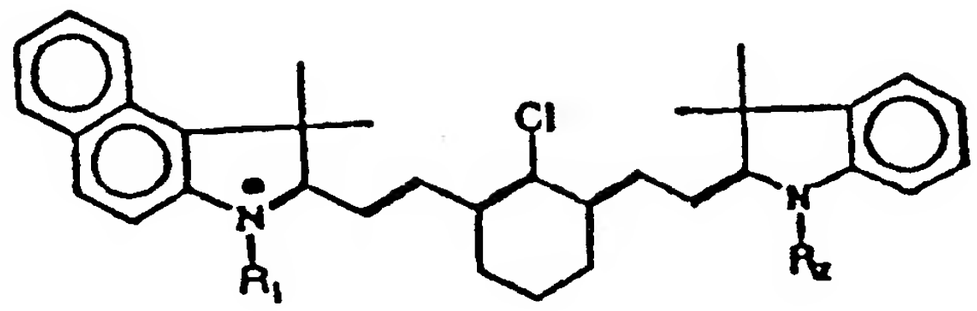

Cl
O
OH
Butanol/Benzene (7:3)
-H2O
No catalyst
Cl
N
R1
4-substituted pheno
DMF/0-5 C
N2or H2
X

CHART 3



Butanol/Benzene (7:3)



4-substituted
phenol/N_aH
DMF 0-5 C

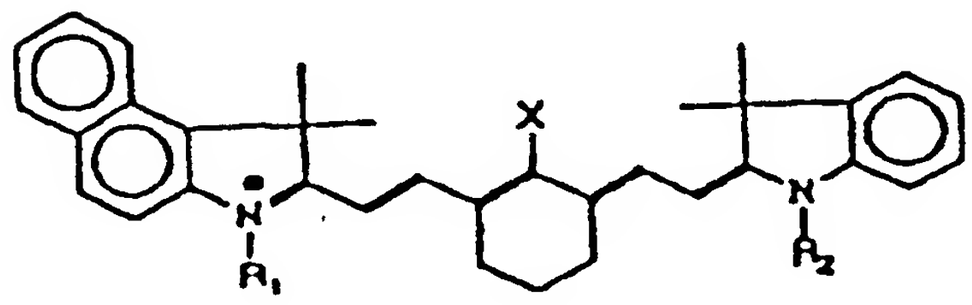


CHART 4

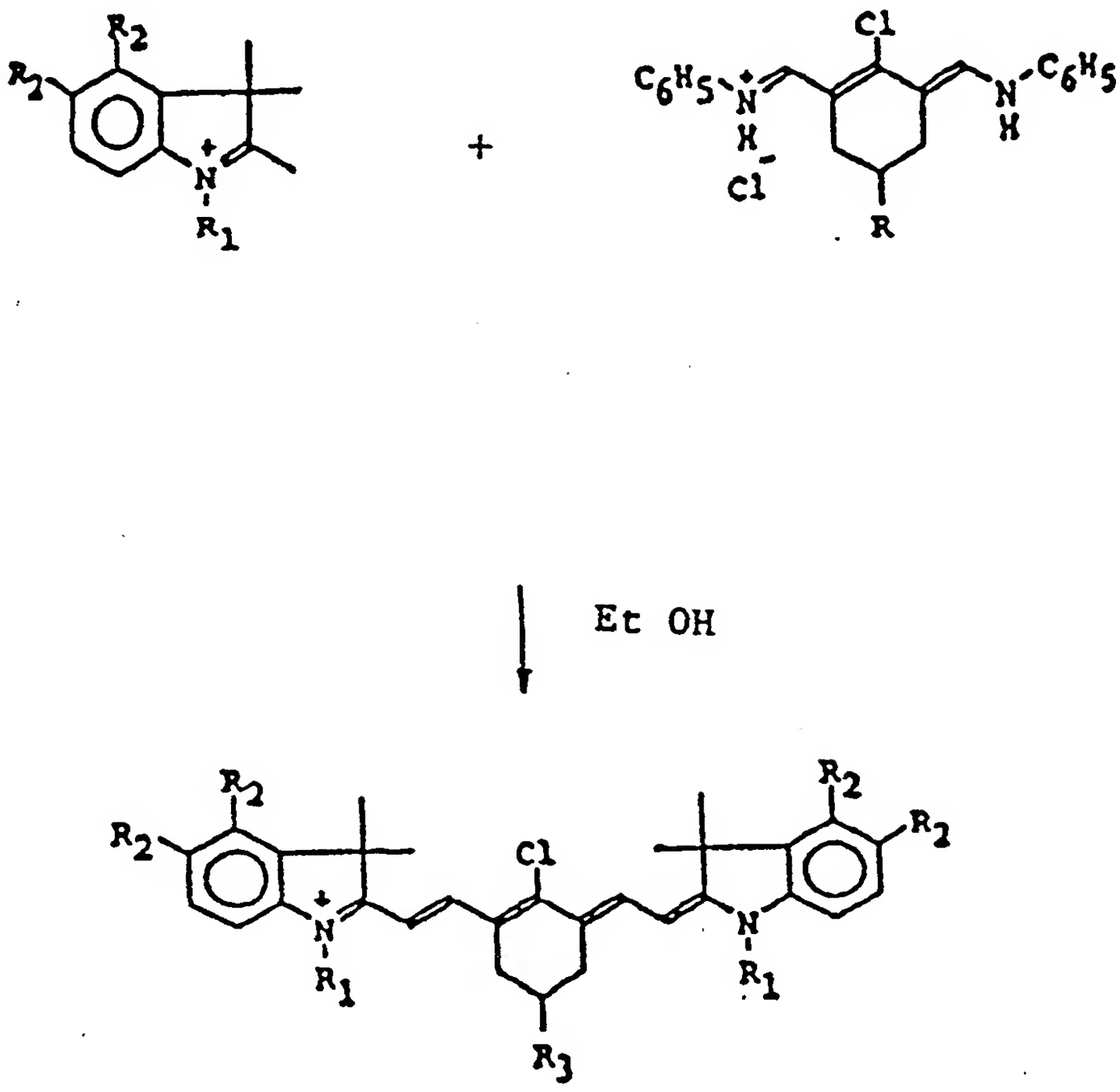
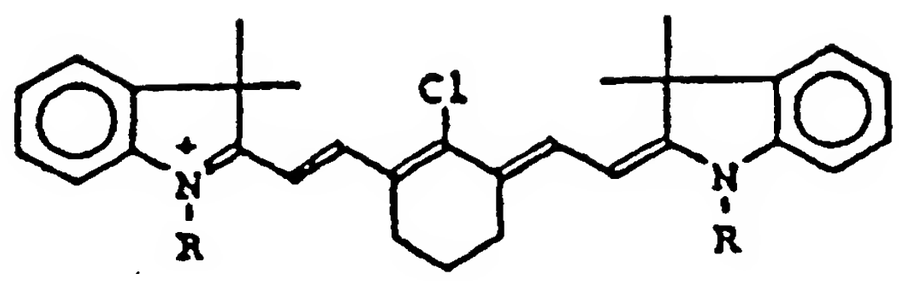
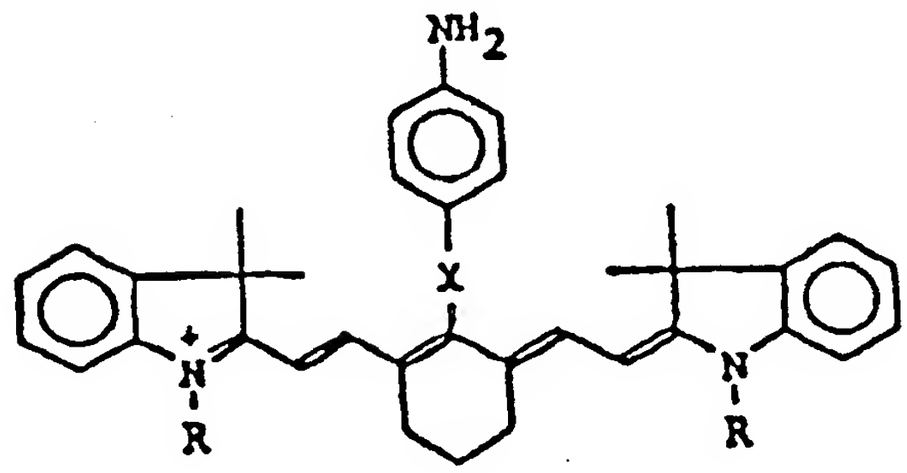


CHART 5





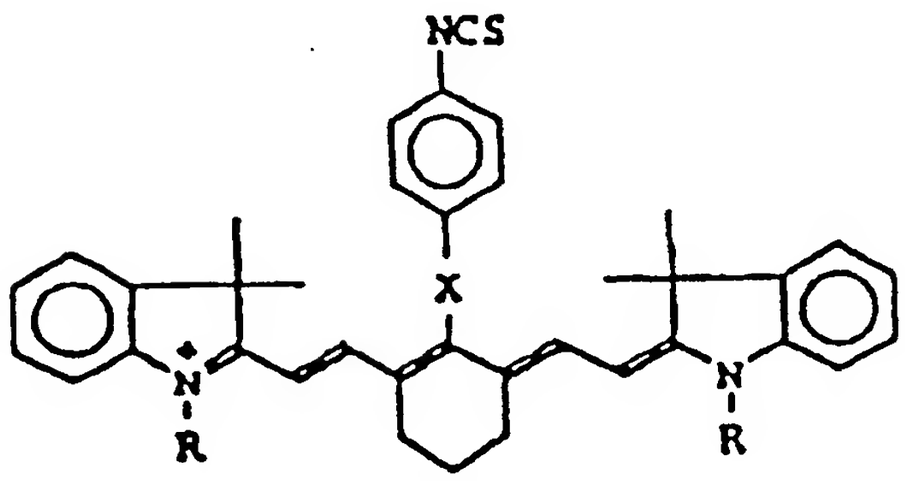
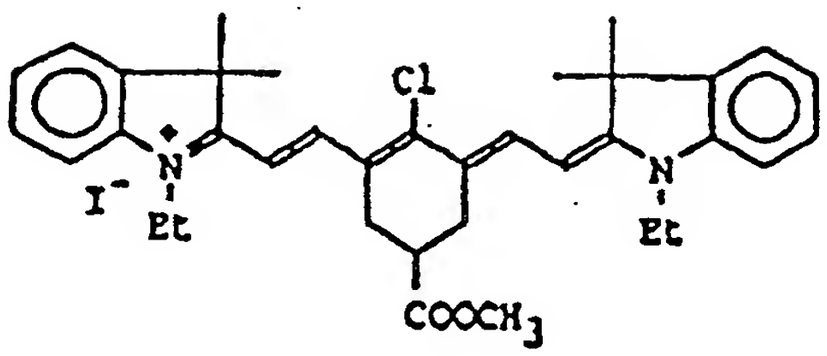
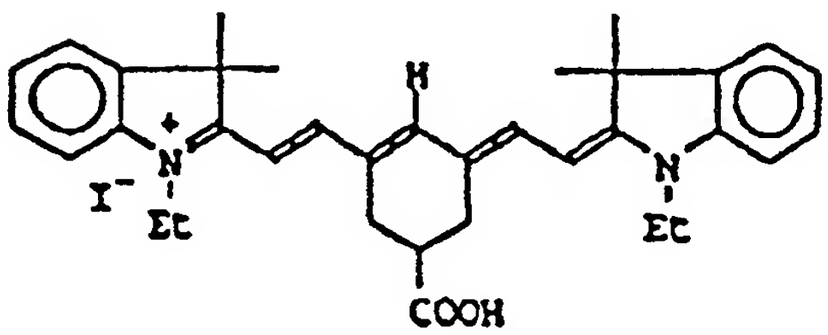
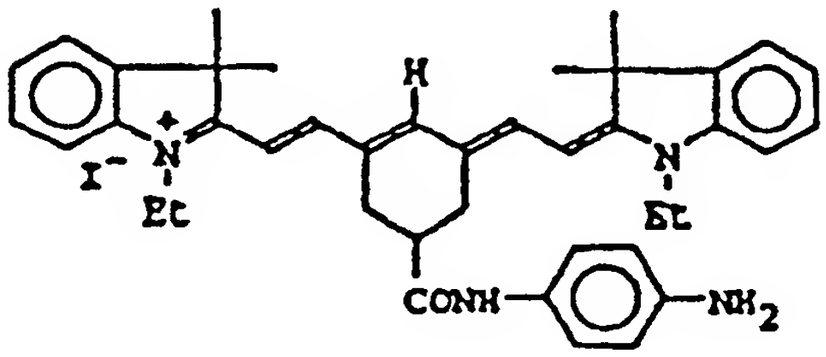


CHART 6







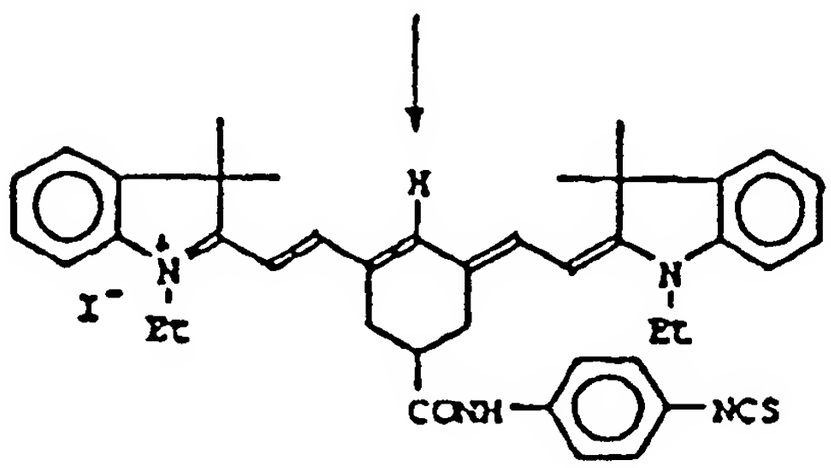
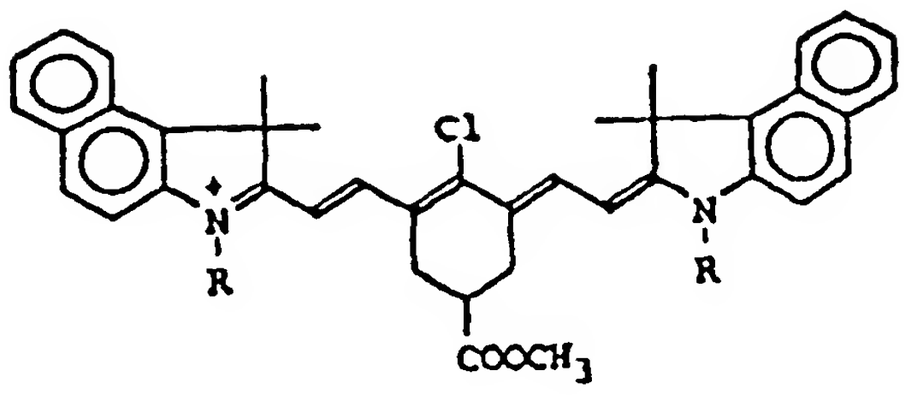
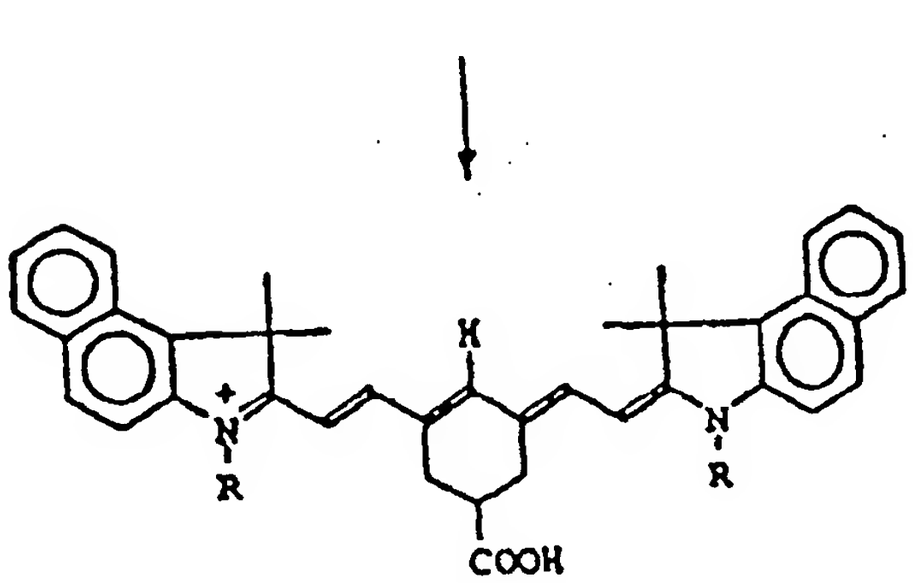
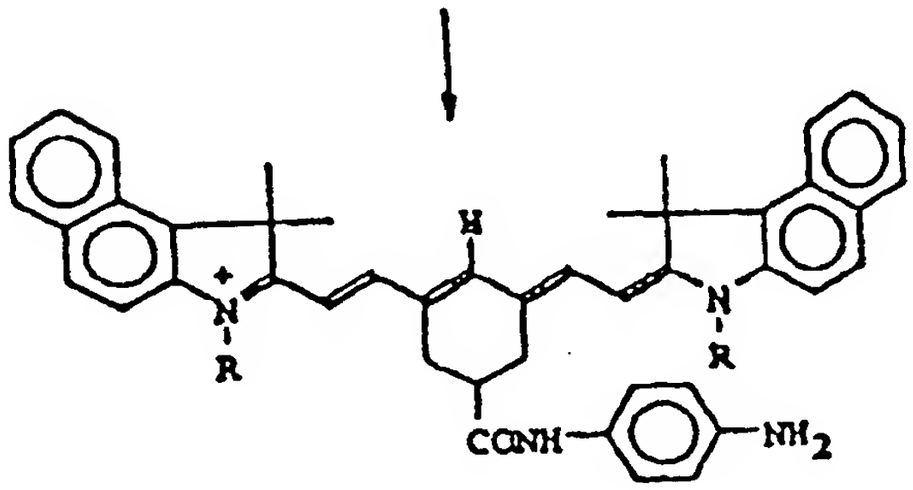
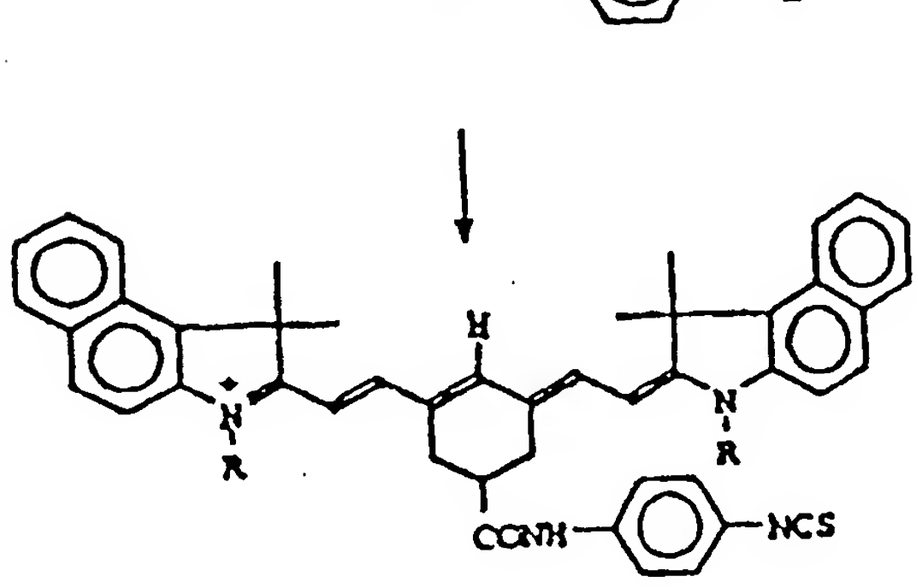


CHART 7









aforementioned United States patent application 07/570,503 filed August 21, 1990; United States patent application 07/078,279 filed July 27, 1987; and United States patent 4,729,947, all of which are entitled DNA SEQUENCING and which were filed by Middendorf et al. on March 29, 1984.

In FIG. 2, there is shown a sectional view of the remote station 122A taken through section lines 2-2 of FIG. 1 having an electrophoresis section 140, a scanning section 142, an electrophoresis power supply 144,

a system power supply section 144A, an analog board 146 and a digital board 148. The electrophoresis section 140 is positioned near the front of the cabinet and a portion of it is adapted to be scanned by the scanning section 142 in cooperation with circuitry on the analog board 146 and the digital board 148. All of the apparatus are electrically connected to the power supply section 144A for such operation.

To separate different DNA fragments into bands, the electrophoresis section 140 includes a gel sandwich 150, an upper buffer assembly 152, support assembly 154, and a lower buffer assembly 151 positioned to enclose the bottom of the gel sandwich 150. In the embodiment of FIG. 2, the gel sandwich 150 is held substantially vertically and its temperature is controlled during operation. Bands are separated by applying voltage to the upper buffer assembly 152 and lower buffer assembly 151 and scanned by the scanning section 142.

To support the gel sandwich 150, the support assembly 154 includes a pair of upper side brackets and lower side brackets 160 and 162 (only one of each pair being shown in FIG. 2), a temperature control heating plate 164, and a plastic spacer, shown at 166A-166C, in FIG. 2. The entire structure is supported on the support assembly 154 which mounts the upper and lower side brackets 160 and 162.

The upper and lower side brackets 160 and 162 are shaped to receive the gel sandwich 150 and hold it in place in juxtaposition with the scanning section 142. The spacer as shown as 166A-166C space the temperature control heating plate 164 from an apparatus support plate 168 and maintain it at a constant selected temperature above ambient temperature. In the preferred embodiment, the temperature is maintained at 50 degrees Centigrade and should be maintained in a range of 30 degrees to 80 degrees.

The scanning section 142 includes a laser diode assembly (not shown in FIG. 2), a microscope assembly 172, a photodiode section 174 and a scanner mounting section 176. The laser diode assembly (not shown in FIG. 2) is positioned at an angle to an opening in the heating plate 164 so that light impinges on the gel sandwich 150 to cause fluorescence with minimum reflection back through the microscope assembly 172.

To receive the fluorescent light, the microscope assembly 172 is focused on the gel sandwich 150 and transmits fluorescent light emitted therefrom into the photodiode section 174 which converts it to electrical signals for transmission to and processing by the analog and digital boards 146 and 148 which may provide further analysis of data. The scanning section 142 moves along a slot in the apparatus support plate 168 which is mounted to the scanner mounting section 176 during this operation in order to scan across the columns in the gel sandwich 150.

The scanner mounting section 176 includes a mounting plate 180, a bearing 182, a stepping motor 184, a slidable support 186 and a belt and pully arrangement 185, 188, 188A. The mounting plate 180 is bolted to the apparatus support plate 168 and supports an elongated bearing plate 182, a stepping motor 184 and two pulleys 188 and 188A. The elongated bearing plate 182 extends the length of the gel sandwich 150.

To permit motion of the laser diode assembly (not shown) and microscope assembly 172 with respect to the gel sandwich 150, the slidable support 186 supports the microscope assembly 172 and diode assembly and slidably rests upon the bearing plate 182. An output shaft 183 of the stepping motor 184 drives a pulley 188B through pulley 188, belt 185, and pulley 188A and the pulley 188B drives a belt (not shown) that is clamped to the slidable support 186 to move it the length of the gel sandwich 150 during scanning by the laser diode and microscope assembly 172 which rest upon it. The stepping motor 184 under the control of circuitry in the digital board 148 moves the pulley 188B to move the belt (not shown) and thus cause scanning across the gel sandwich 150.

As shown in this view, the electrophoresis power supply 144 is electrically connected to buffer in the upper buffer assembly 152 through an electrical connector 194 and to the lower buffer assembly 151 through a connector not shown in FIG. 2.

The upper buffer assembly 152 includes walls 197 forming a container to hold a buffer solution 195 and a cover 199 formed with a lip to fit over the walls 197 from the top and containing a downwardly extending flat member spaced away from the side walls and holding a conductor 211. The conductor 211 is electrically connected to the source of power through connector 194 which is mounted to the top of the cover 199 to permit electrical energization of the buffer solution 195.

The bottom buffer assembly 151 includes enclosed walls 201 defining a container for holding a buffer solution 203 and a cap 205 closing the container 201 and having a downwardly extending portion 213 extending into the buffer 203 for supporting a conductor 207 for applying energy to the bottom buffer solution 203. The gel sandwich 150 extends downwardly into the buffer solution 203 and upwardly into the buffer solution 195 to permit the electrical contact for electrophoresis.

In FIG. 3, there is shown a sectional view taken through lines 3-3 of FIG. 1 showing the electrophoresis section 140, the scanning section 142 and the electrophoresis power supply section 144 mounted together to illustrate from a top view the arrangement of the apparatus support plate 168, the heater plate 164, the

gel sandwich 150, a microscope assembly 172 and a photodiode assembly 174. The heater plate 164 and apparatus support plate 168 have slots running in a horizontal direction orthogonal to the lanes of DNA in the electrophoresis section 140 sized to receive the ends of a laser diode assembly 170 and the microscope section 172 for scanning thereof.

To cooperate with the separation and scanning of DNA bands, the gel sandwich 150 includes a front glass plate 200, a gel section 202 and a rear glass plate 204 mounted in contact with the heater plate 164 and having a section exposed for scanning by the laser diode assembly 170 and the microscope assembly 172. The rear glass plate 204 contacts the heater plate 164 and is separated from the front plate 200 by the gel section 202 within which DNA separation takes place. The front and rear glass plates may be any type of glass but are preferably soda lime which has low fluorescence in the infrared and near infrared regions and is prepared by a process that provides optically flat surfaces without grinding.

To transmit light to the gel sandwich 150, the laser diode assembly 170 includes a housing 210, a focusing lens 212, a narrow band pass filter 214, a collimating lens 216 and a laser diode 218. The laser diode 218 emits infrared or near infrared light which is collimated by the laser collimating lens 216 and filtered through the narrow band pass infrared filter 214. This light is focused by the focusing lens 212 onto the gel sandwich 150. Preferably, the point of focus on the gel section 202 of the gel sandwich 150 lies along or near the central longitudinal axis of the microscope section 172 and the photodiode section 174.

The thickness of the glass plates and the gel, the position of the laser and sensor and their angle of incidence are chosen, taking into consideration the refractive index of the gel and glass so that the light from the laser is absorbed by a maximum number of markers for one channel. The light from the laser is not directly reflected back because the angle of incidence to normal is equal to the Brewster's angle at the first interface and is such as to impinge on the markers with full intensity after refraction but not be reflected by subsequent layers of the gel sandwich 150 into the sensor and the sensor views a large number of markers that fluoresce in a line of sight of substantial concentration.

To maintain temperature control over the laser diode, the housing 210: (a) is coupled to a heat sink through a thermal electric cooler 220, and (b) encloses the focusing lens 212, narrow band pass filter 214, collimating lens 216 and laser diode 218; and (c) accommodates the electrical leads for the diode. To receive and focus light emitted by fluorescent markers from the gel section 202 in response to the light from the laser diode assembly 170, the microscope assembly 172 includes a collection lens 230, a housing 232 and a coupling section 234. The microscope assembly 172 is adapted to be positioned with its longitudinal axis centered on the collection lens 230 and aligned with the photodiode section 174 to which it is connected by the coupling section 234. For this purpose, the housing 232 includes a central passageway in which are located one or more optical filters with a band pass matching the emission fluorescence of the marked DNA strands along its longitudinal axis from the axis of the collection lens 230 to the coupling section 234 which transmits light to the photodiode section 174. With this arrangement, the collection lens 230 receives light from the fluorescent material within the gel section 202 and collimates the collected light for optical filtering and then transmission to the photodiode assembly 174.

To generate electrical signals representing the detected fluorescence, the photodiode assembly 174 includes a housing 240 having within it, as the principal elements of the light sensors, an inlet window 242, a focusing lens 244, a sapphire window 246 and an avalanche photodiode 248. To support the avalanche photodiode 248, a detector mounting plate 250 is mounted within the housing 240 to support a plate upon which the avalanche photodiode 248 is mounted. The inlet window 242 fits within the coupling section 234 to receive light along the longitudinal axis of the photodiode assembly 174 from the microscope assembly 172.

Within the housing 240 of the photodiode assembly 174, the sapphire window 246 and avalanche photodiode 248 are aligned along the common axis of the microscope assembly 172 and the photodiode assembly 174 and focuses light transmitted by the microscope assembly 172 onto a small spot on the avalanche photodiode 248 for conversion to electrical signals. A thermoelectric cooler 252 utilizing the Peltier effect is mounted adjacent to the detector mounting plate 250 to maintain a relatively cool temperature suitable for proper operation of the avalanche photodiode 248.

The lower buffer assembly 151 (FIG. 2) includes outer walls 201 and a bottom wall forming a compartment for buffer solution which encloses the bottom of the gel sandwich 150.

As best shown in this view, the stepping motor 184 rotates the belt 185 to turn the pulley 188A, which, in turn, rotates pulley 188B. The pulley 188B includes a belt 177 extending between it and an idler pulley 179 and attached at one location to the slideable support 186 to move the scanning microscope and laser lengthwise along the gel sandwich 150 for scanning purposes. The motor 184, by moving the carriage back and forth, accomplishes scanning of the gel sandwich 150.

In FIG. 4, there is shown a fragmentary perspective view of the gel sandwich 150 and the upper buffer assembly 152 mounted to each other showing the outer glass plate 200 cut away from the rear glass plate 204 to expose the gel section 202 to buffer solution within the upper buffer assembly 152. With this arrangement, samples may be pipetted between the glass plates 200 and 204 and moved downwardly by electrophoresis beyond the upper buffer assembly 152 and through the gel sandwich 150 to the bottom buffer (not shown in FIG. 4).

In FIG. 5, there is shown a broken away view of the gel sandwich 150 illustrating the upper buffer assembly 152 and the lower buffer assembly 151 connected to it at each end. As shown in this view, the cover 199 includes a connecting post 214 which receives the conductor 211 for connection to the downwardly extending portion of the cover 199 into the buffer compartment. Between the glass plates 200 and 204 (FIG. 4) of the gel sandwich 150, are a plurality of downwardly extending recesses 221 in the gel section 202 (FIG. 4) between the plates. DNA sample is pipetted into these recesses to form channels for electrophoresing to the lower buffer assembly 151.

To form an electical connection through the gel sandwich 150 from the upper buffer assembly 152 to the lower buffer assembly 151, a conducting post 216 is connected to the cover 205 of the lower buffer assembly 151 for receiving the conductor 207 which extends downwardly to the downwardly extended plate 213 and into the buffer solution.

In FIG. 6, there is shown a block diagram of the circuitry used to control the remote station 122A of the embodiment of FIG. 2 having a control, correlation and readout section 250, the scanner drive 176, the motor assembly 184 for moving the scanner drive 176, and the sensing configuration 252. The sensing configuration 252 includes the laser assembly 170 and the sensor assembly 174 which receives signals, removes some noise, and transmits the signals for display and read out in the control, correlation and read out section 250, while the scanner drive 176 and motor for the scanner drive 184 receive signals from the control, correlation and read out section 250 to control the motion of the sensor back and forth across the gel sandwich. This overall configuration is not part of the invention of this application except insofar as it cooperates with the sensing configuration 252 to scan the DNA and determine its sequence in accordance with the embodiments of FIGS 1-5.

To drive the sensor 174 from position to position, the motor assembly 184 includes a stepper motor 254 and a motor driver 256. The motor driver 256 receives signals from the control correlation and read-out section 250 and actuates the stepper motor 254 to drive the scanner drive 176. The scanner drive 176 is mechanically coupled to a stepping motor 254 through a belt and pulley arrangement for movement back and forth to sense the electrophoresis channels on the gel sandwich 150 (FIG. 3). The stepping motor 254 and driver circuitry 256 are conventional and not themselves part of the invention.

The control, correlation and read out system 250 includes a computer which may be any standard microprocessor 260, a television display or cathode ray tube display 262 and a printer 264 for displaying and printing the results of the scans.

To sense data, the sensing configuration 252 includes, in addition to the laser 170 and the sensor 174, a chopper circuit 270, a sensor power supply 272, a preamplifier 274, a lock-in amplifier 276, a 6-pole filter 278, a 12-bit analogue digital converter interface circuit 280 and a laser power supply 282.

The sensor 174 receives light from the laser 170 after it impinges upon the gel sandwich 150 (FIG. 3) and transmits the signals through preamplifier 274 to the lock-in amplifier 276. The sensor receives signals from the sensor power supply 272. The chopper circuit 270 provides pulses at synchronized frequencies to the lock-in amplifier 276.

The laser 170 receives power from the power supply 282 which is controlled by the chopper circuit 270 so that the signal from the laser is in synchronism with the signal applied to the lock-in amplifier 276 so that the output from the lock-in amplifier 276 to the 6-pole filter 278 discriminates against unwanted signal frequencies. This signal is converted to a digital signal in the 12-bit analogue to digital converter 280 which serves as an interface to the computer 260.

With this arrangement, the scanning rate may be set to discriminate against noise and the synchronized demodulation from the chopper control further reduces noise, particularly discriminating against the natural fluorescense of the glass in the gel sandwich 150 (FIGS. 2 and 3).

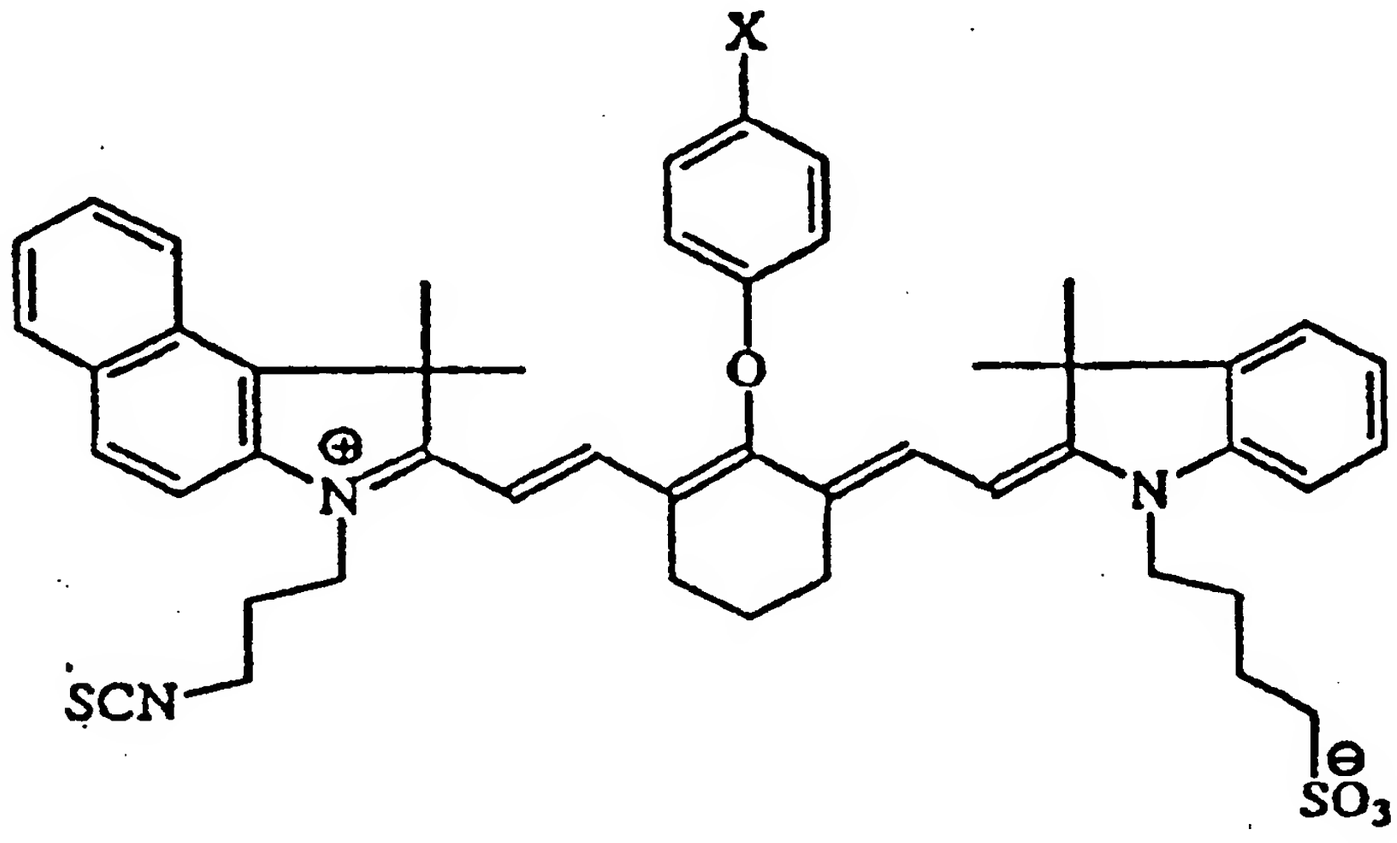
From the above summary, it can be understood that the techniques for the sequencing of fluorescence labeled DNA of this invention have several advantages, such as: (1) because the dyes have their emission spectra in the infrared or near infrared light spectrum, small inexpensive infrared diode lasers may be used; and (2) they are characterized by relatively low noise.

Although a preferred embodiment of the invention has been described with some particularity, many modifications and variations are possible in the preferred embodiment within the light of the above description. Accordingly, within the scope of the appended claims, the invention may be practiced other

than as specifically described.

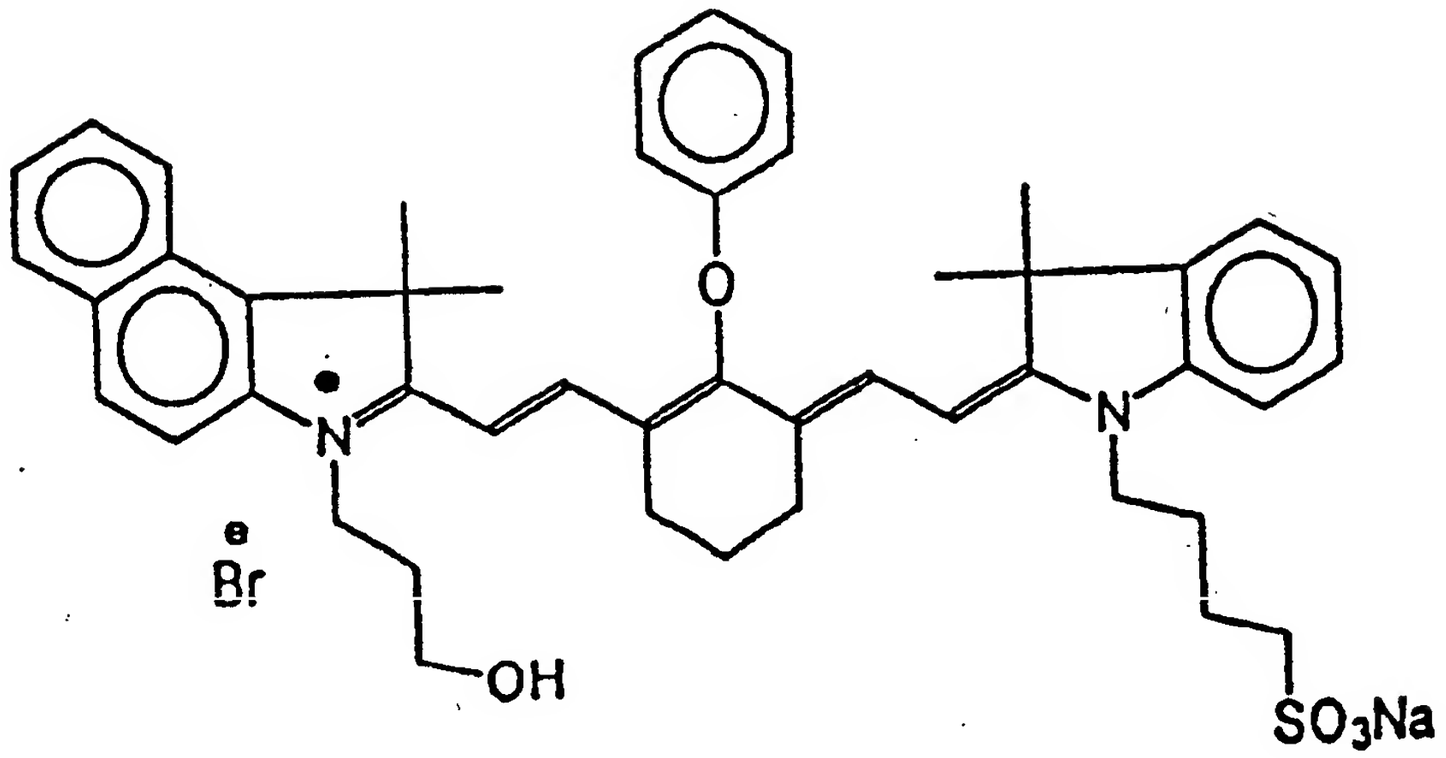
Claims

1. A method of identifying strands of DNA comprising the steps of: marking the strands with fluorescent labels, irradiating the strands and detecting the light emitted from the fluorescent strands, characterized in that the labels emit light in a region of wavelengths including at least one wavelength within the infrared and near infrared region wherein the fluorescent label includes a chromophore having one of the formulas:

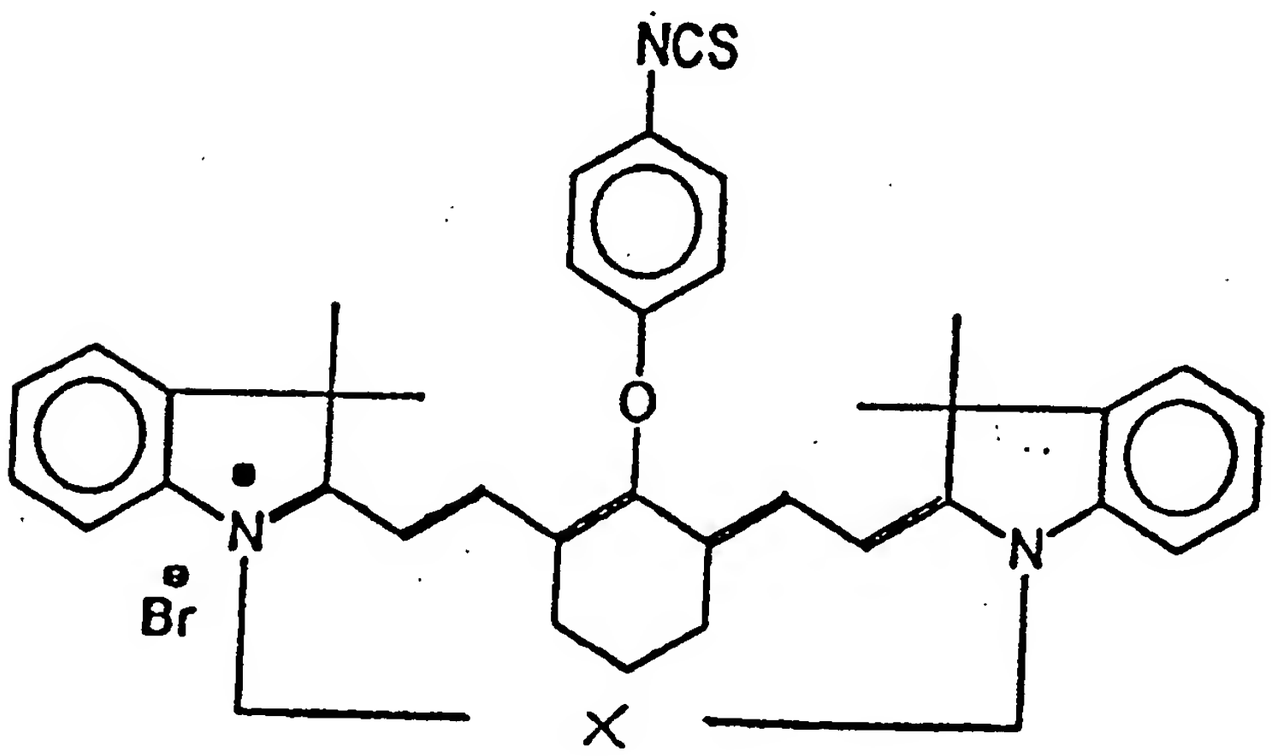


BNSDOCID: <EP 0670374A1 I >

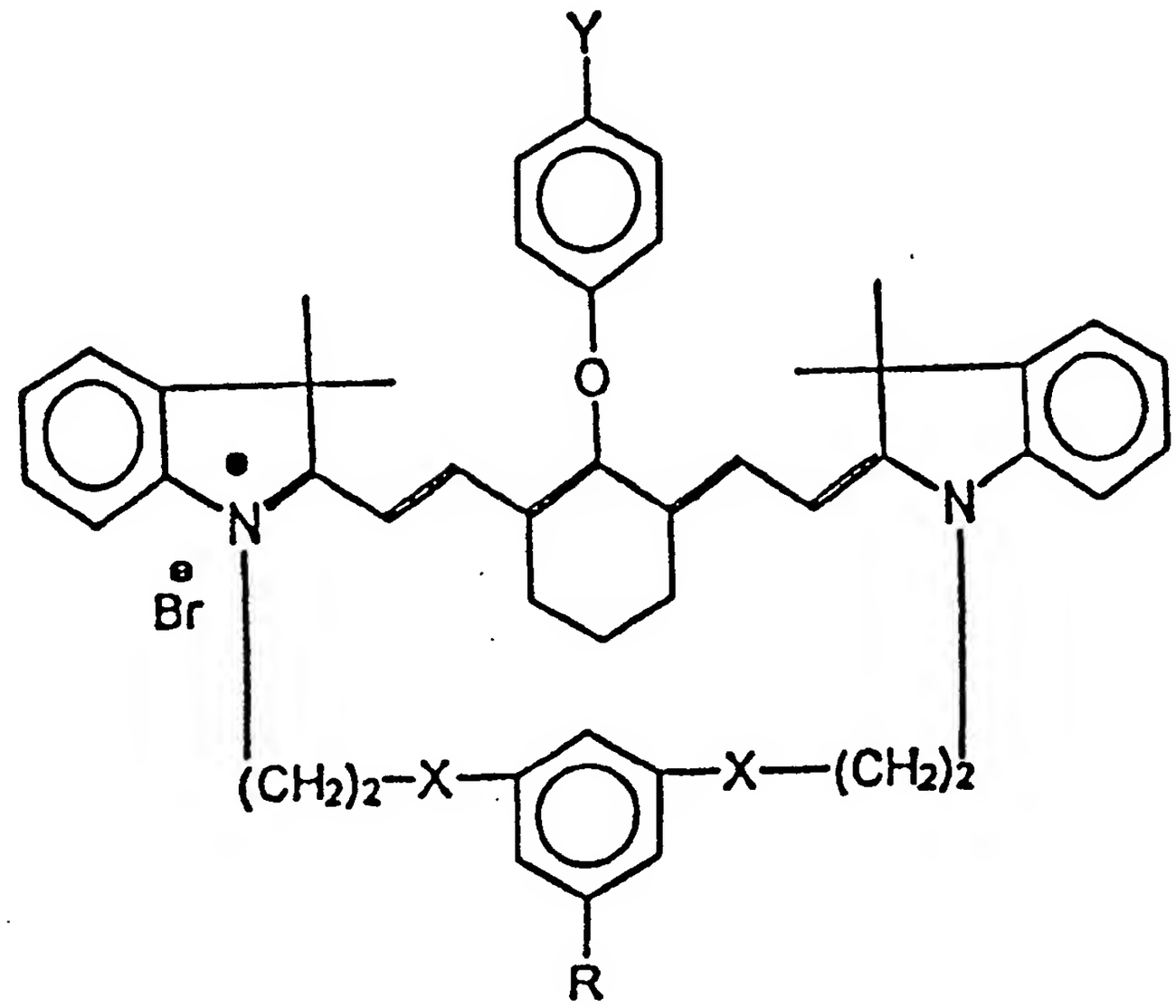
or



or



where X is $(CH_2)_n$; n = 4-10 or X is -CH_2 - CH_2 - 0 - CH_2 - CH_2 - 0 - CH_2 - CH_2 -; or

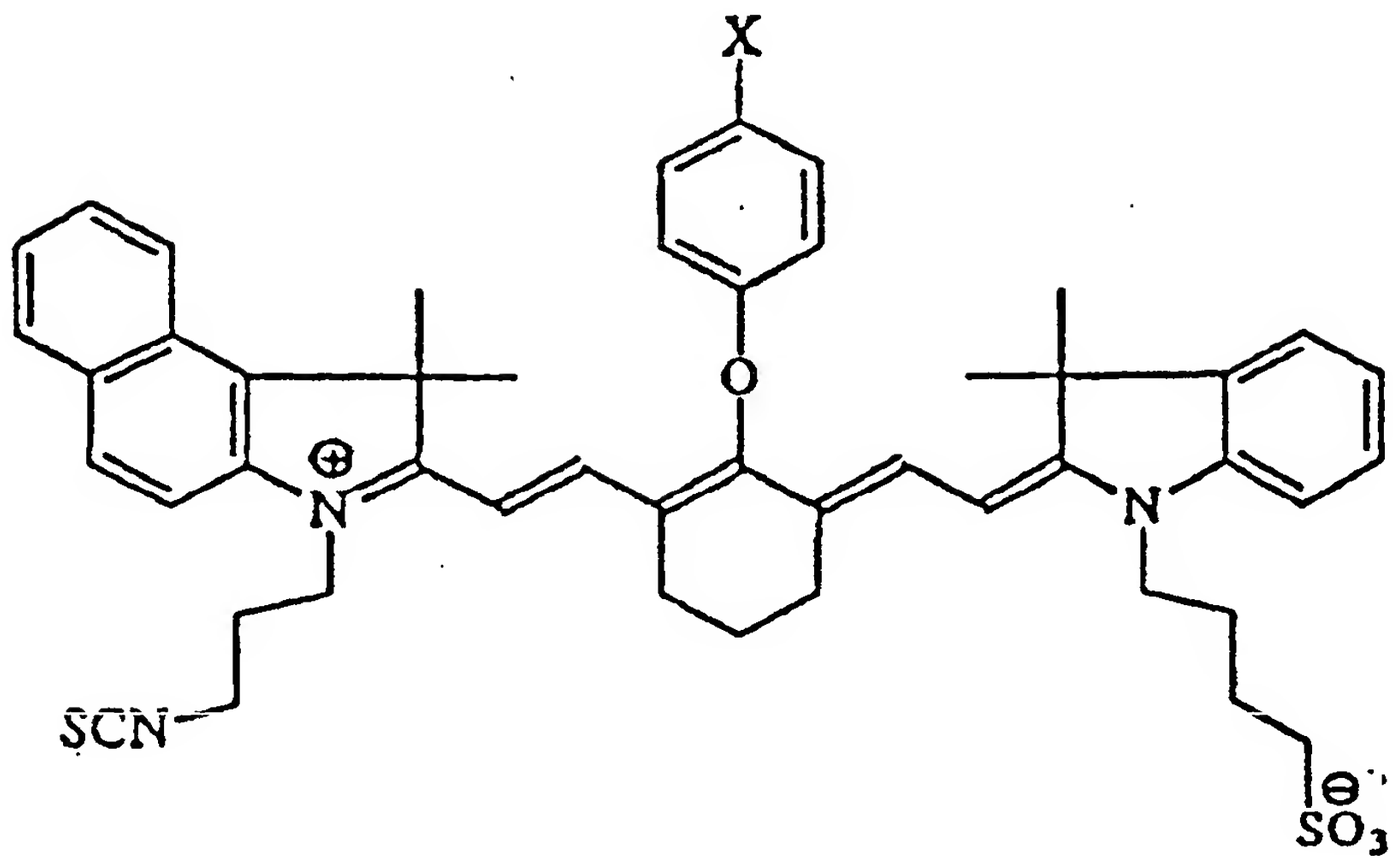


wherein X consists of one of O or NH, Y consists of one of NCS or H; and R consists of one of H, NCS, CH_2OH, CH_2NCS, COOH; the light used to irradiate the strands has a wavelength within one of the infrared and near infrared regions.

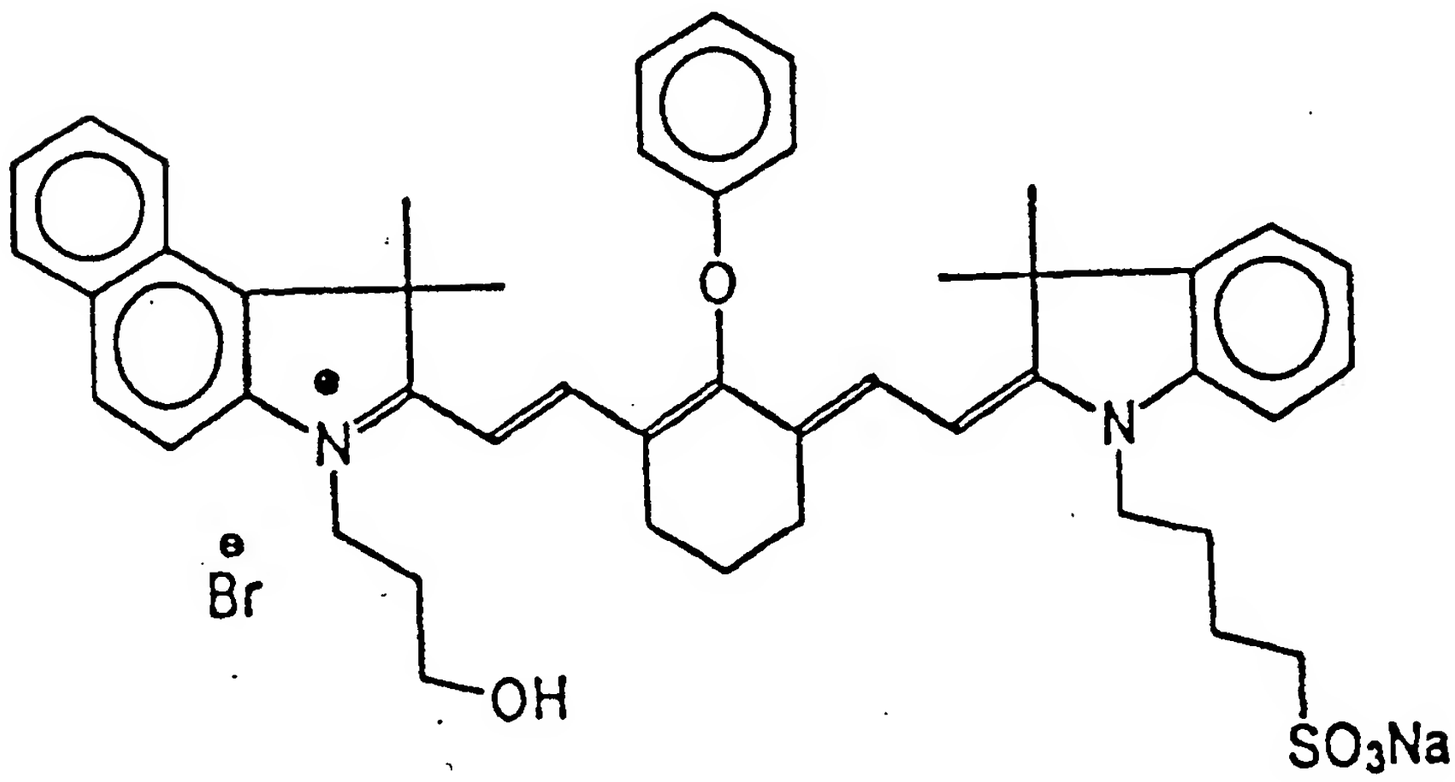
2. A method of identifying strands of DNA in accordance with claim 1 characterized by scanning with an infrared laser diode light source (218).

3. A method according to either claim 1 or claim 2 characterized by: applying DNA samples marked with fluorescent dye having absorbance and fluorescense maxima at near infrared or infrared wavelengths when combined with the DNA at a plurality of locations in a gel electrophoresis slab (202) for electrophoresing in a plurality of channels through the gel electrophoresis slab (202); establishing electrical potential across said gel electrophoresis slab (202) wherein DNA samples are resolved in accordance with the size of DNA fragments in said gel electrophoresis slab (202) into fluorescently marked DNA bands; and scanning the separated samples photo-electrically while they are in the slab with a laser diode (218) and a sensor (174) wherein the laser (218) scans with near infrared or infrared scanning light at a scanning light frequency within the absorbance spectrum of said fluorescently marked DNA samples and sensing light at the emission frequency of the marked DNA.

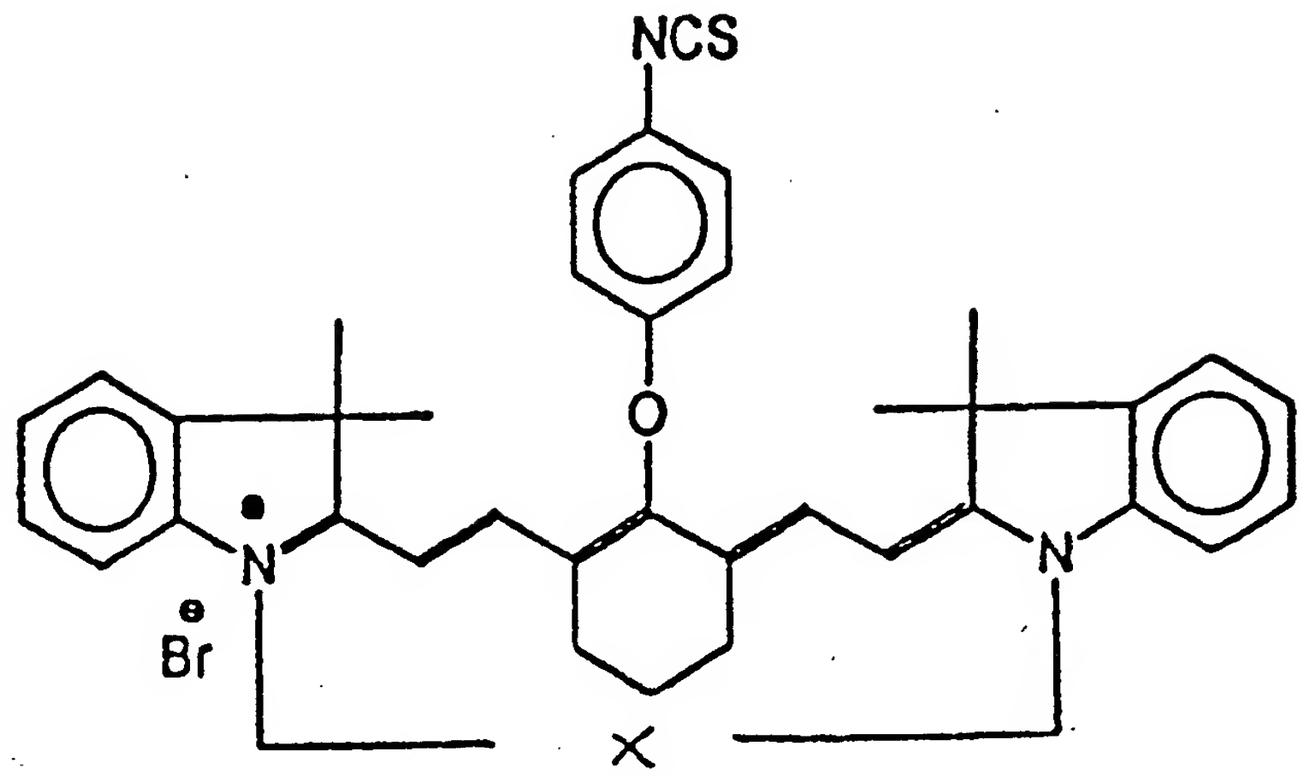
4. A dye of any of the following formulas;



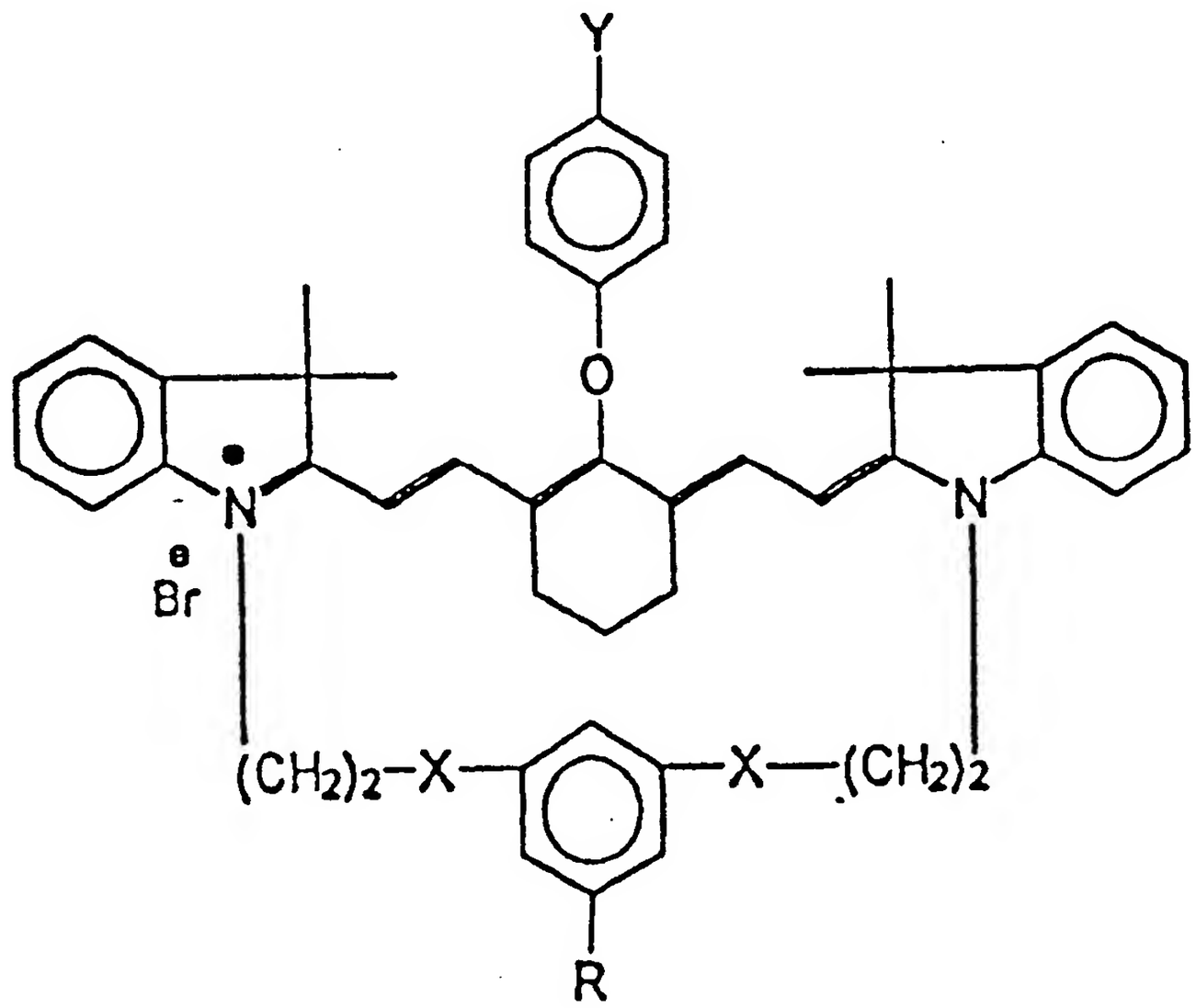
or



or



where X is $(CH_2)_n$; n = 4-10 or X is -CH_2 - CH_2 - 0 - CH_2 - CH_2 - 0 - CH_2 - CH_2 -; or



wherein X is one of O or NH; Y is one of NCS or H; R is one of H, NCS, CH_2OH, CH_2NCH or COOH.

FIG. 1

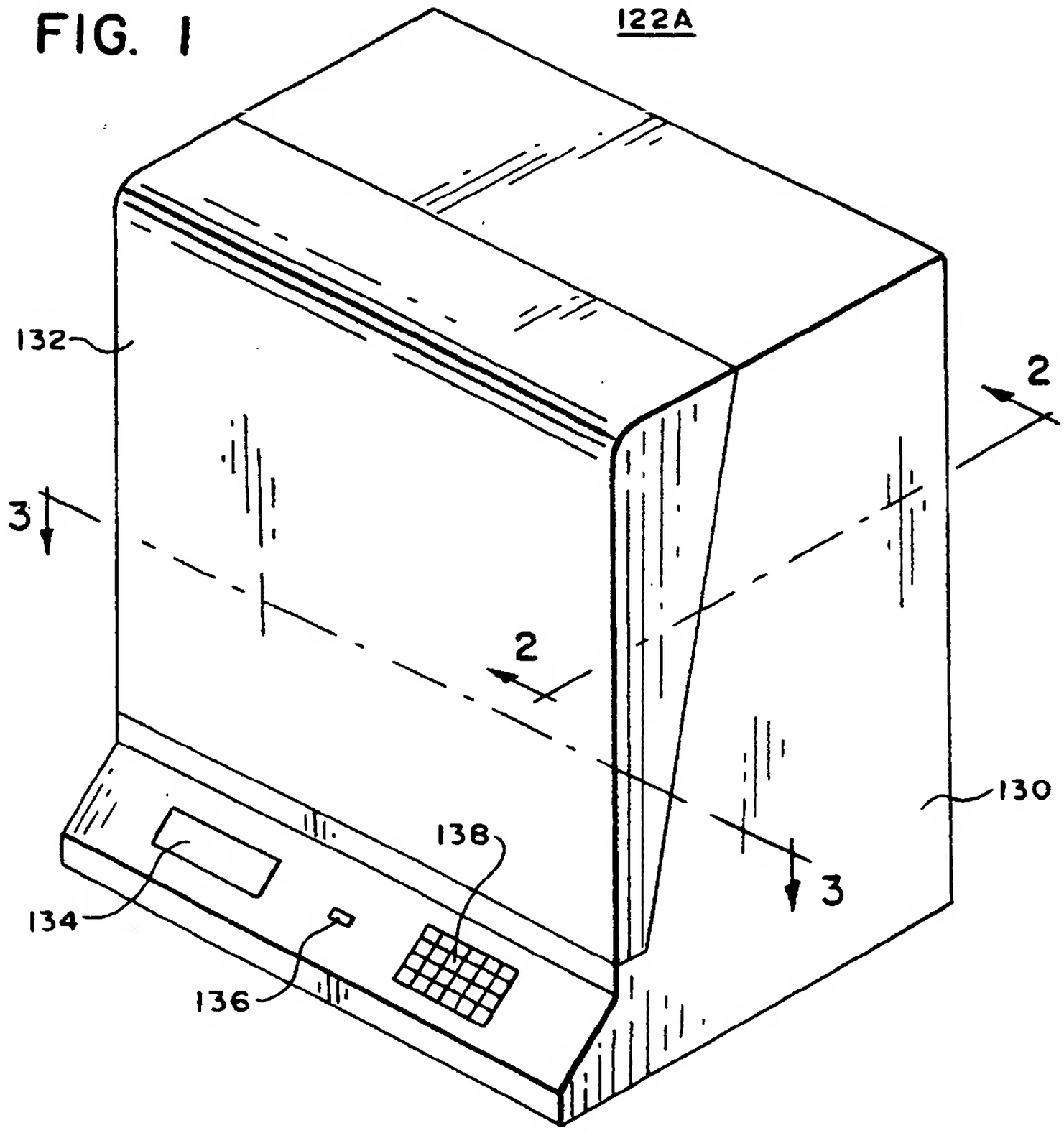


FIG. 2

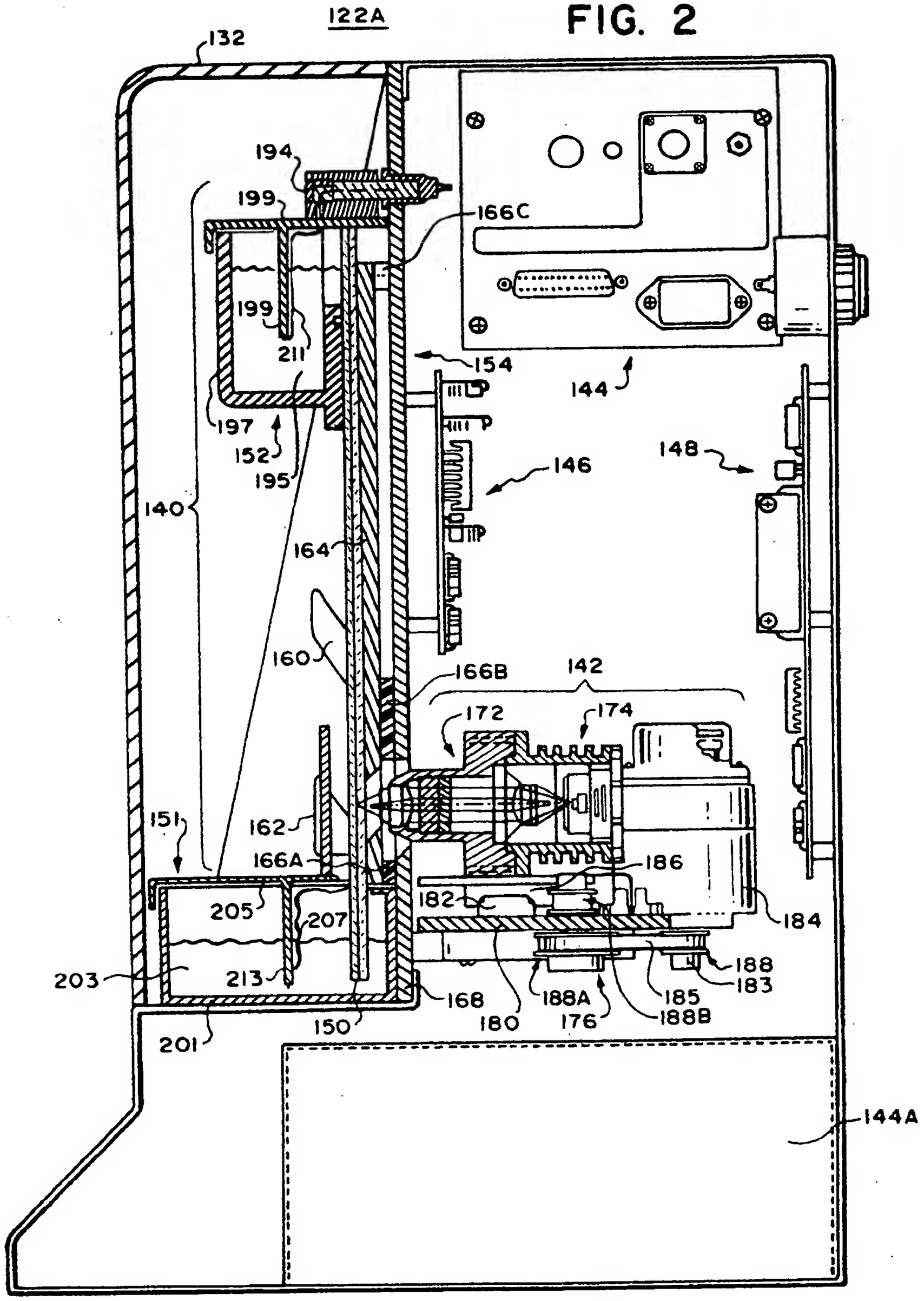

122A
132
194
199
166C
199
211
154
144
197
152
195
148
146
140
164
160
166B
142
172
174
151
162
166A
186
182
184
205
207
188
183
203
213
168
188A
185
180
176
188B
150
201
144A

FIG. 3

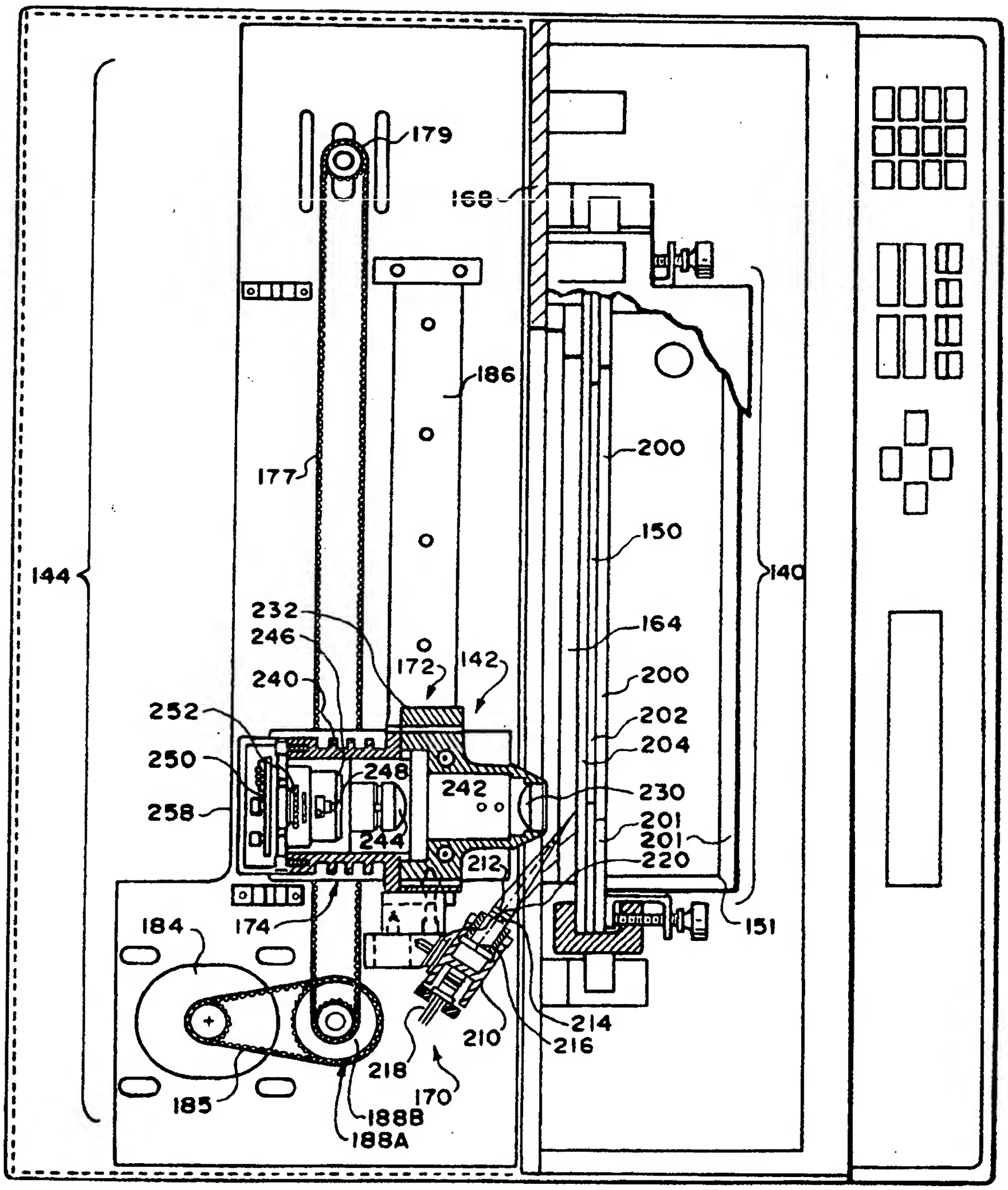

179
168
186
177
144
140
200
150
232
246
240
172
142
164
200
202
204
252
250
248
242
230
258
244
212
201
201
220
184
174
151
214
216
210
218
170
185
188B
188A

FIG. 4

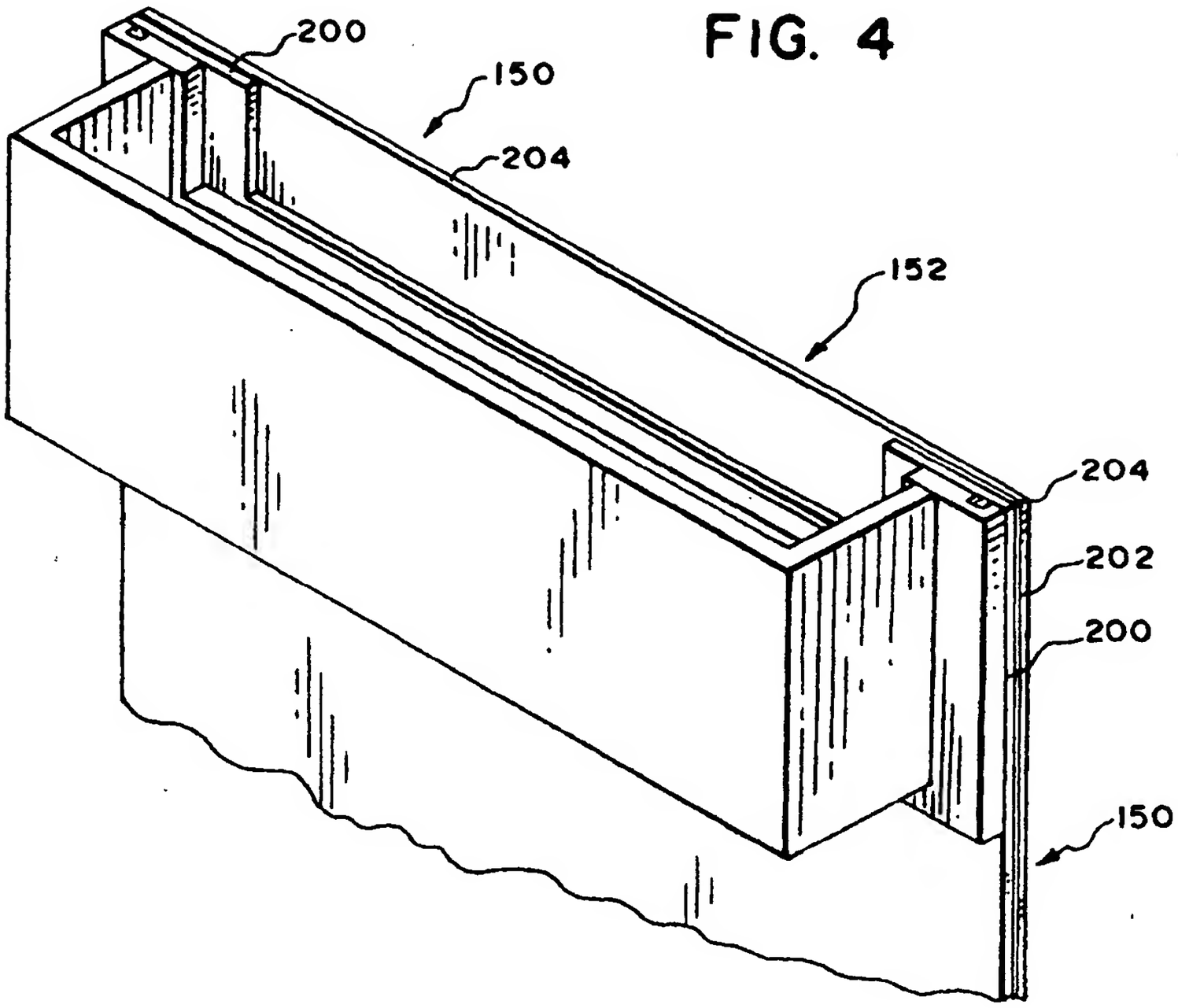
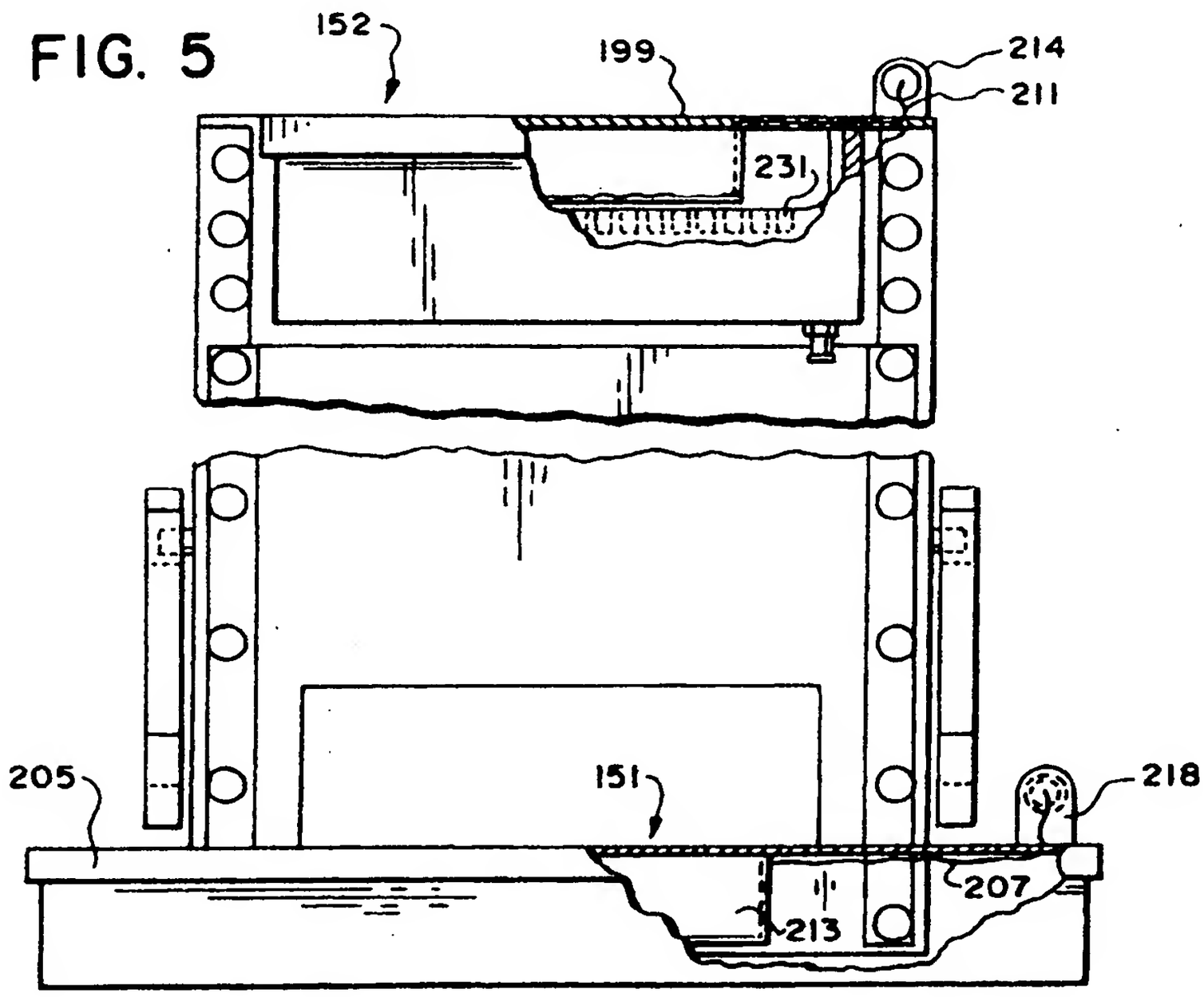


FIG. 5



BNSDOCID: <EP 0670374A1 I >

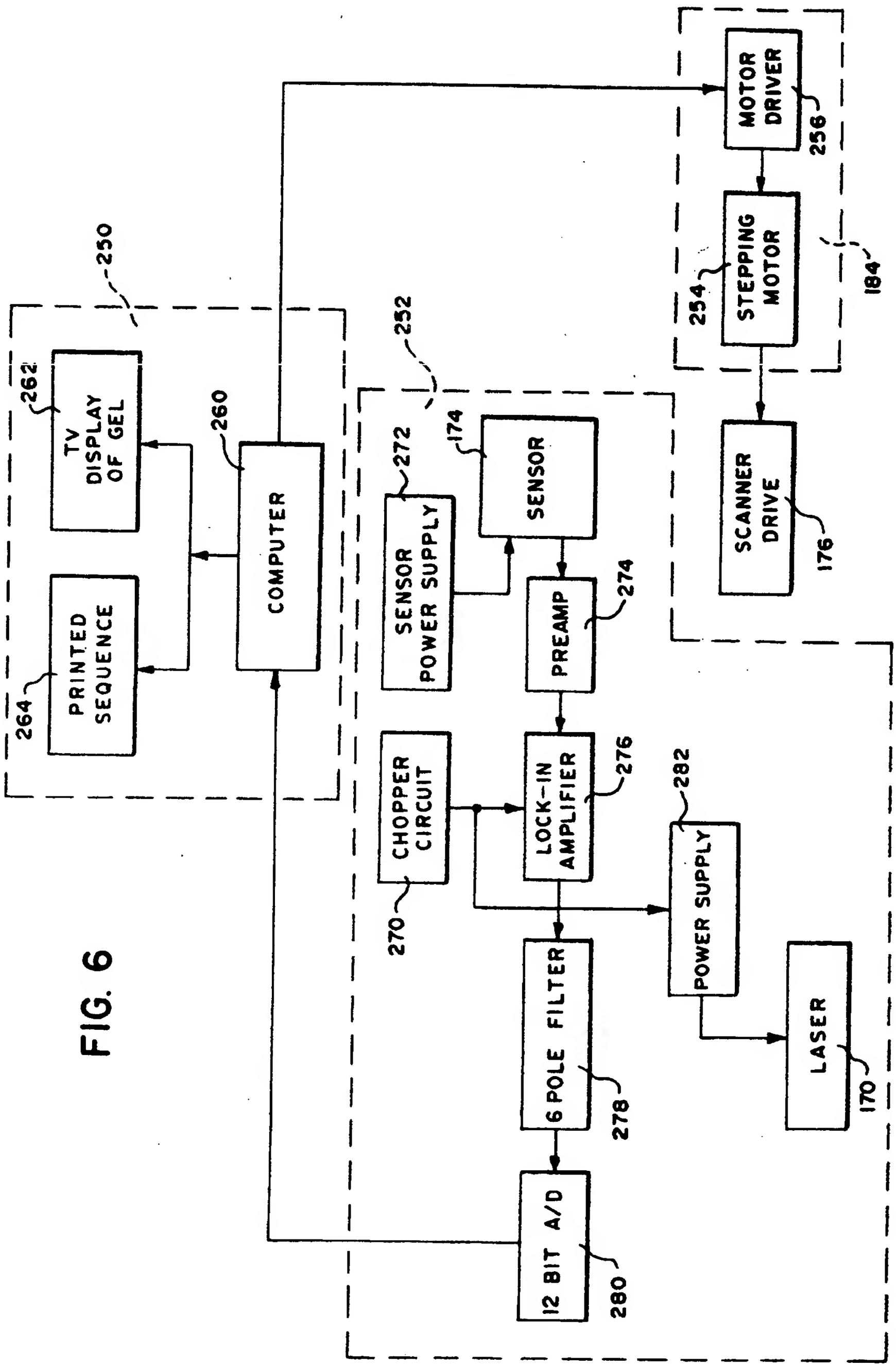

250
252
262
TV DISPLAY OF GEL
260
COMPUTER
264
PRINTED SEQUENCE
MOTOR DRIVER
256
254
STEPPING MOTOR
184
SCANNER DRIVE
176
272
SENSOR POWER SUPPLY
174
SENSOR
274
PREAMP
270
CHOPPER CIRCUIT
276
LOCK-IN AMPLIFIER
282
POWER SUPPLY
6 POLE FILTER
278
12 BIT A/D
280
LASER
170

FIG. 6

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 95 25 0047

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Y	EP-A-0 533 302 (LI-COR, INC.) * the whole document * ---	1-4	C12Q1/68 G01N33/52 C07H21/00 G01N27/447
Y	SYNTHETIC COMMUNICATIONS, vol.23, no.21, 1993 pages 3087 - 3094 M.LIPOWSKA ET AL. 'New Near-Infrared Cyanine Dyes For Labelling of Proteins.' * the whole document * ---	1-4	
Y	JOURNAL OF ORGANIC CHEMISTRY., vol.57, no.17, 1992, EASTON US pages 4578 - 4580 L.STREKOWSKI ET AL. 'Substitution Reactions of a Nucleofungal Group in Heptamethine Cyanine Dyes.' -----	1-4	

TECHNICAL FIELDS SEARCHED (Int.Cl.6)

C07H
C12Q
G01N

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	15 March 1995	Scott, J

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

FIG. 1

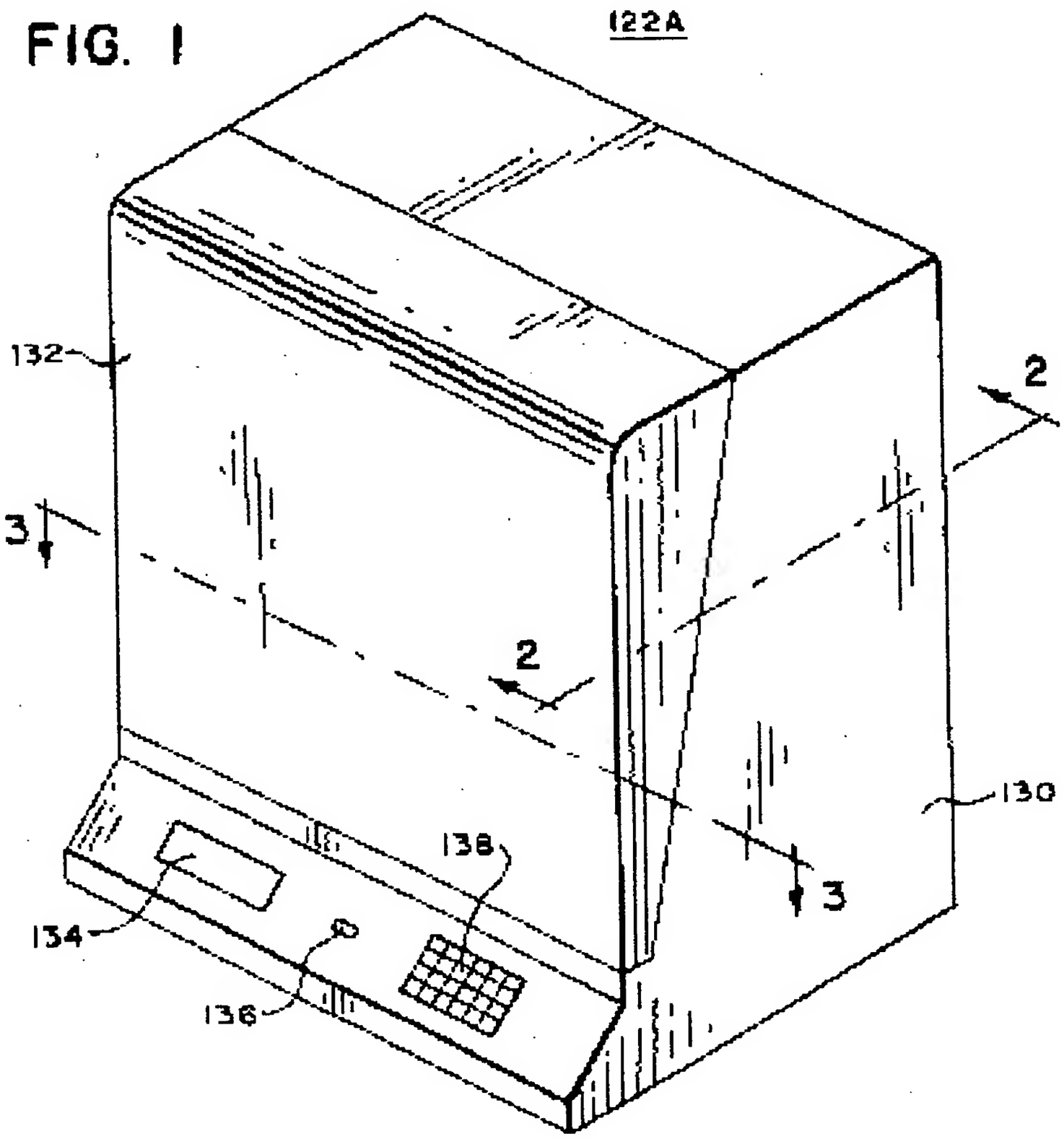


FIG. 2

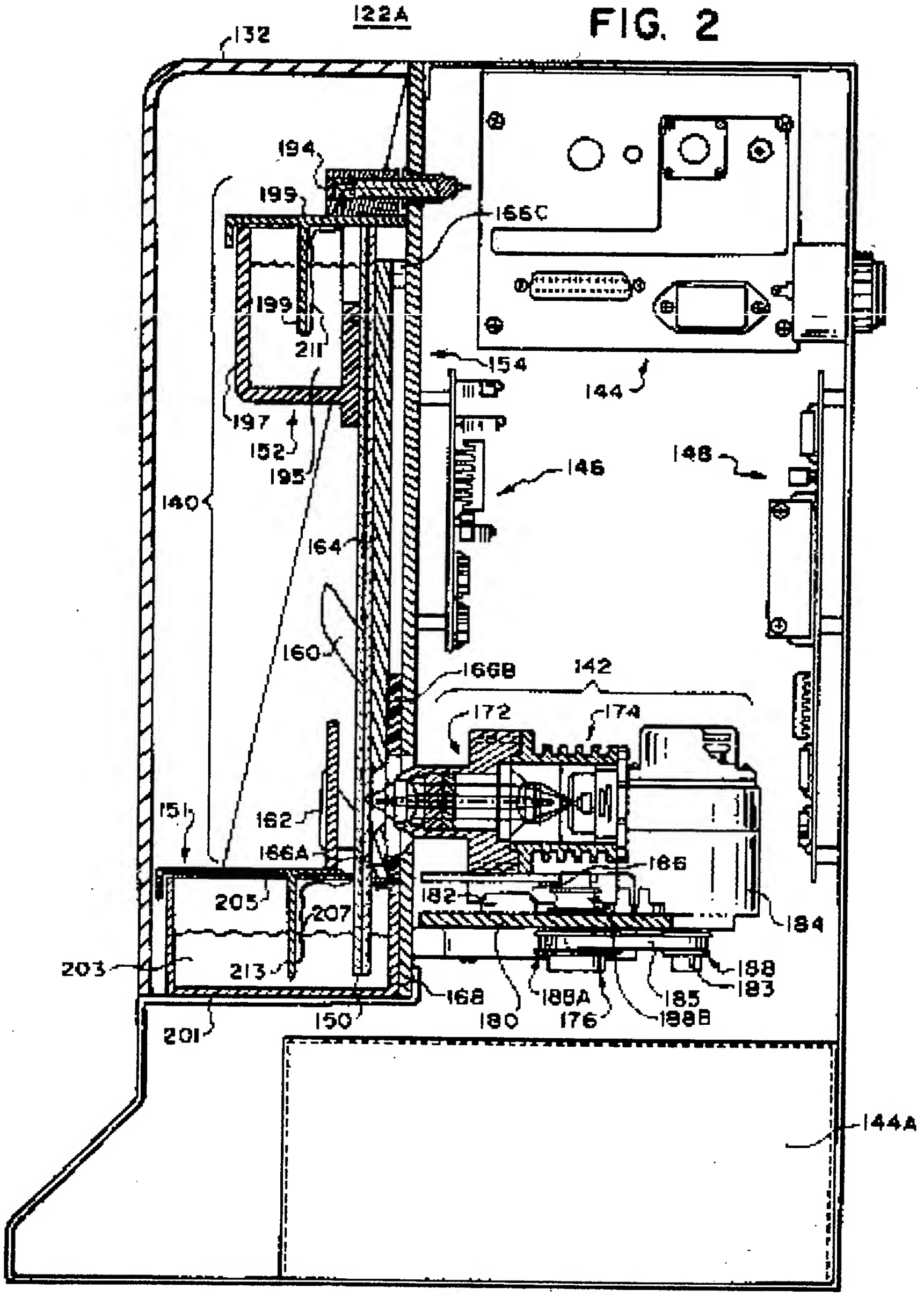

122A
132
194
199
165C
199
211
154
144
197
152
195
148
146
140
164
160
166B
142
172
174
151
162
166A
166
205
207
182
184
203
213
188
183
168
188A
185
180
176
188B
150
201
144A

FIG. 3

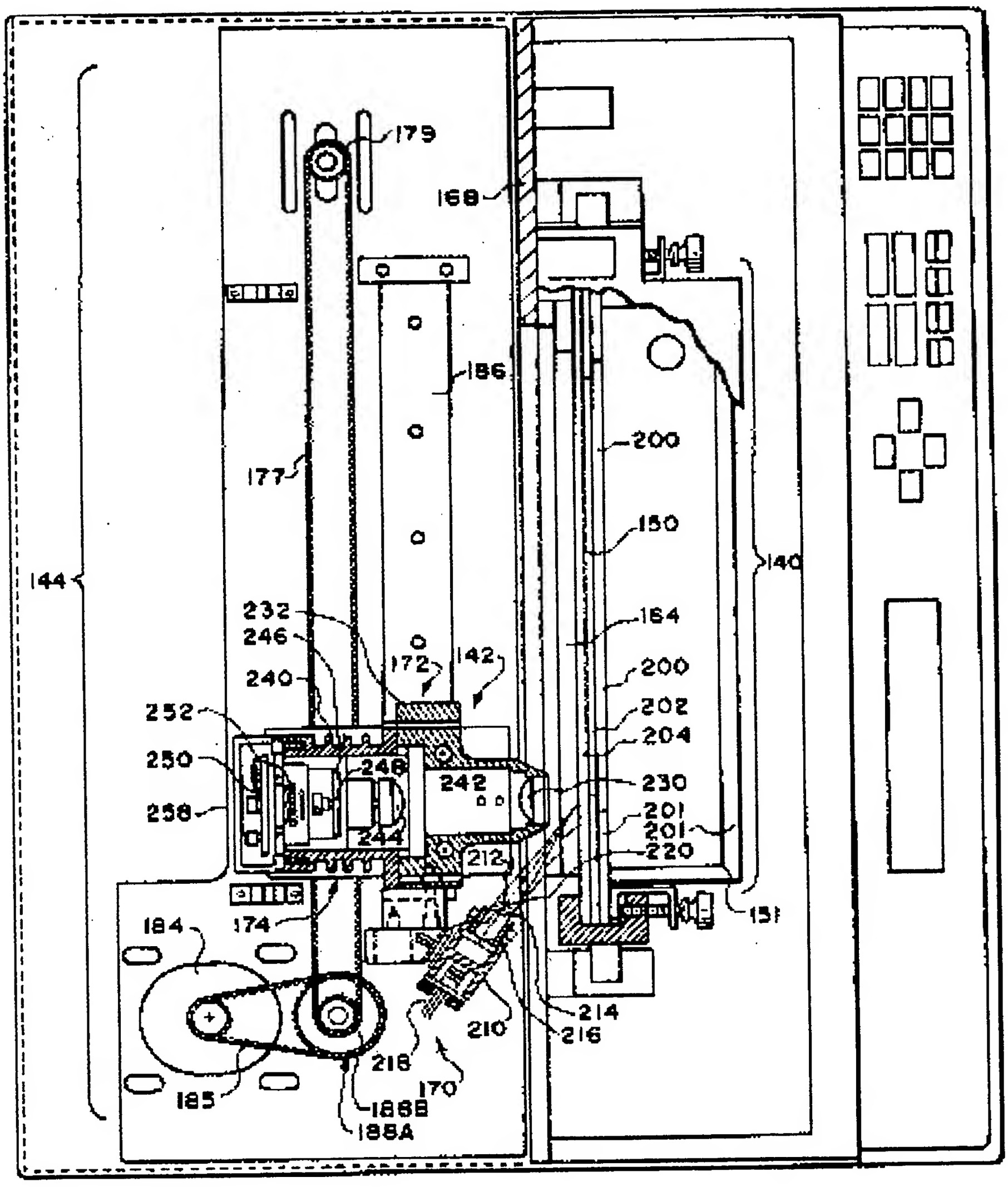

179
168
186
177
200
150
140
144
232
246
240
172
142
164
200
202
204
252
250
246
242
230
201
201
220
258
244
212
151
184
174
214
216
210
218
170
185
188B
188A

FIG. 4

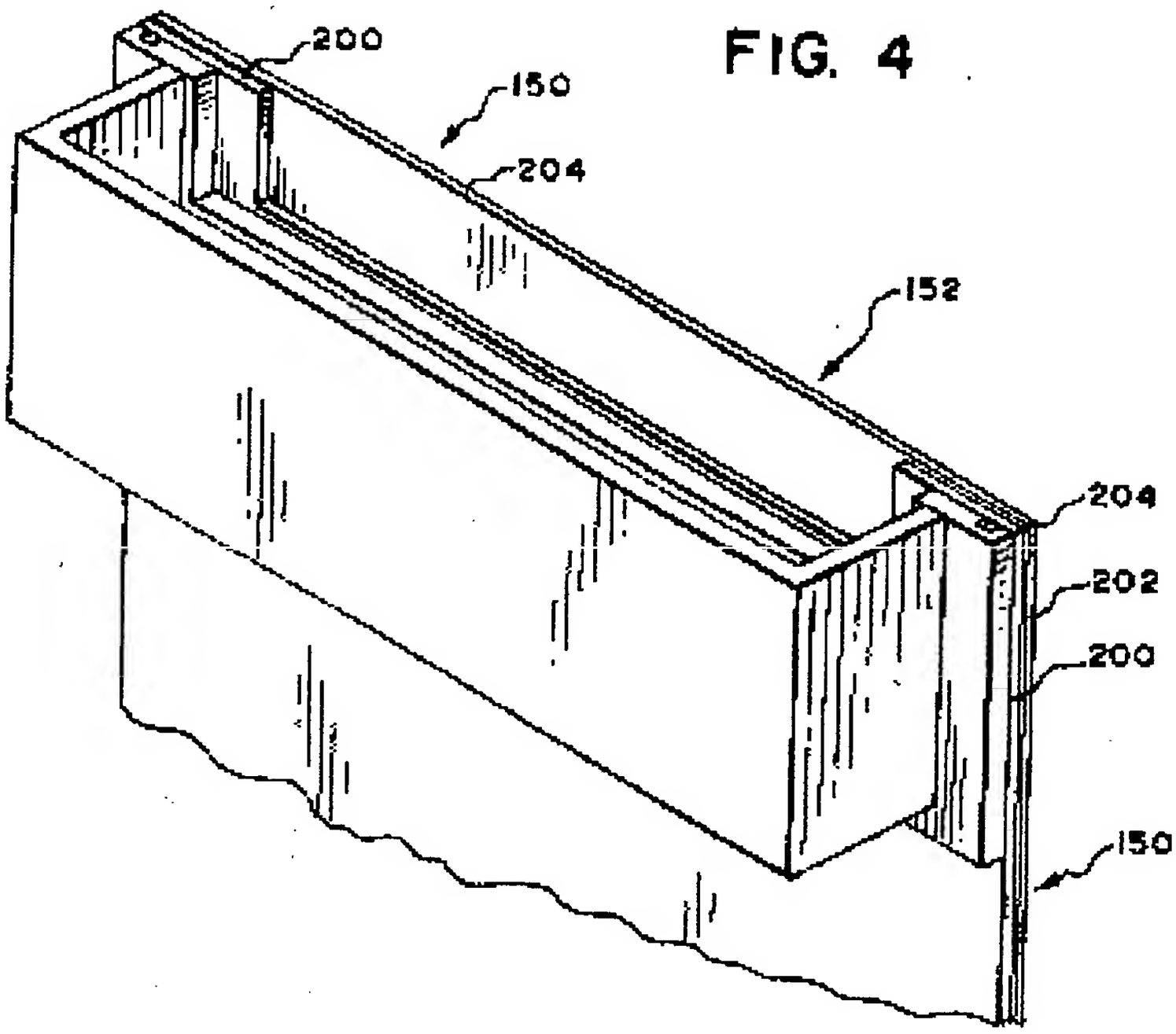
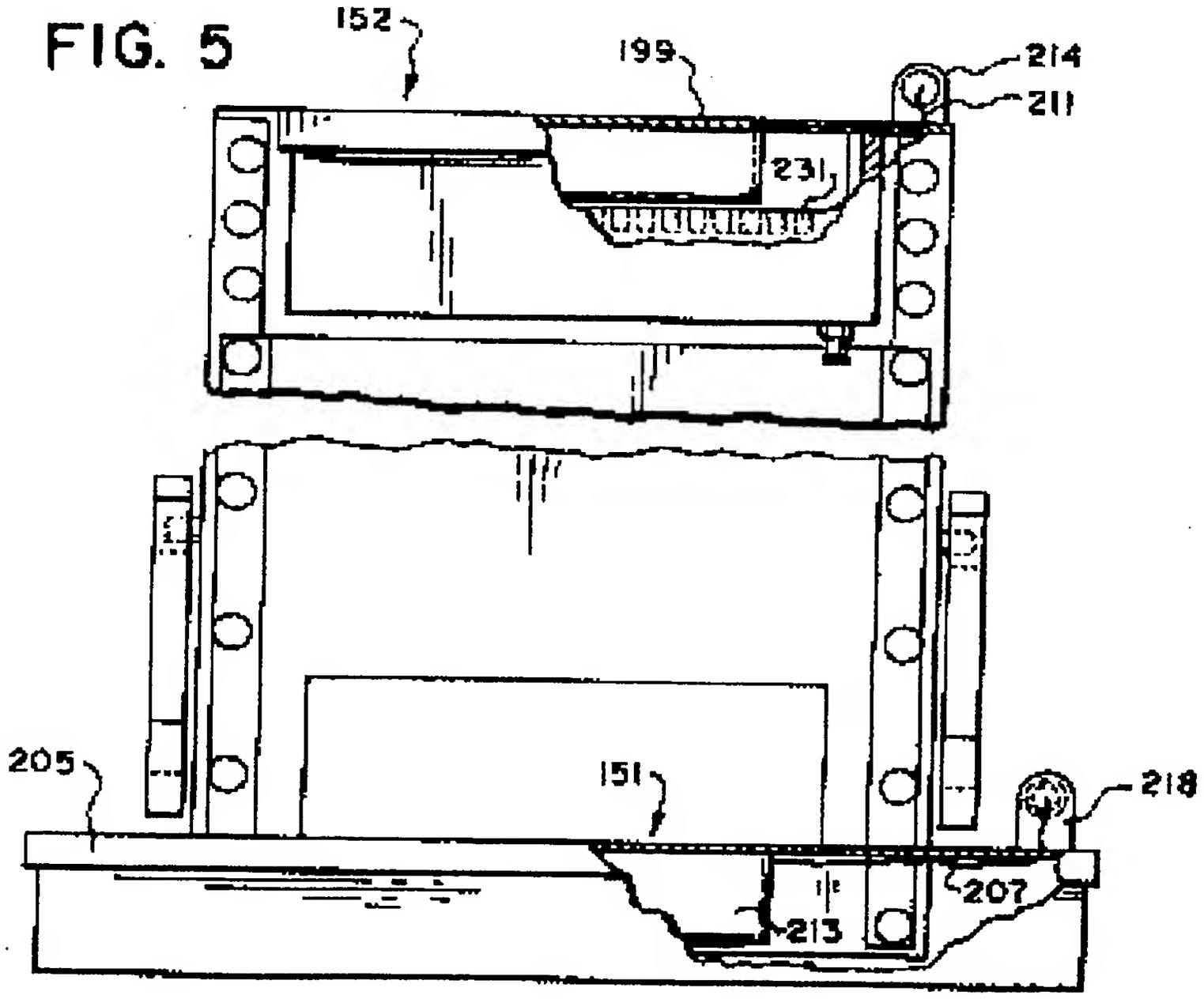

200
150
204
152
204
202
200
150

FIG. 5


152
199
214
211
231
205
151
218
207
213

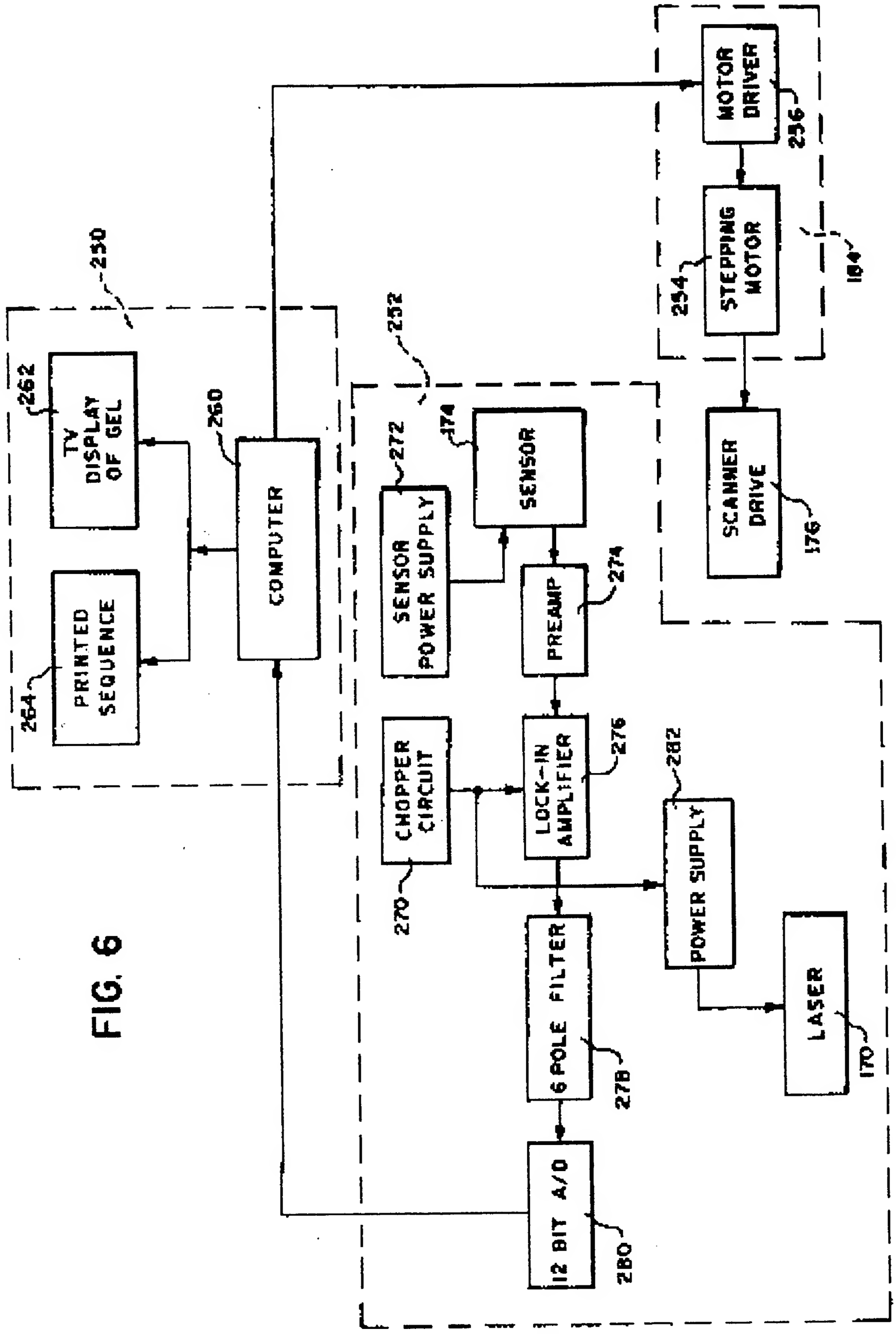

MOTOR DRIVER
236
STEPPING MOTOR
234
184
250
252
262
TV DISPLAY OF GEL
260
COMPUTER
272
174
SENSOR
SENSOR POWER SUPPLY
SCANNER DRIVE
176
264
PRINTED SEQUENCE
CHOPPER CIRCUIT
PREAMP
274
LOCK-IN AMPLIFIER
276
282
270
6 POLE FILTER
278
LASER
170
12 BIT A/D
280

FIG. 6